# Trilateral Impact Assessment Framework for Automation in Road Transportation

Trilateral Impact Assessment Sub-Group for ART

Satu Innamaa, Scott Smith, Yvonne Barnard, Lydia Rainville, Hannah Rakoff, Ryota Horiguchi, Helena Gellerman

April 2018

Version 2.0

| Version | Date | Contribution |
|---|---|---|
| 1.0 | 4 Jan 2017 | Version sent to Trilateral ART WG on 4 Jan 2017 |
| 2.0 | 29 March 2018 | Public version 2.0 |

# Contents

# 1 Introduction

Automated vehicles (AVs) have the potential to transform the world's road transportation system. Benefits could include traffic safety (automobile crashes are a leading cause of accidental deaths), transport network efficiency (most cities experience significant traffic congestion), energy/emissions (oil consumption, air pollution and greenhouse gas emissions are of worldwide concern) and personal mobility (non-drivers may enjoy new mobility options). AVs are being introduced into a complex transportation system. Second order impacts, such as the possibility of increased travel leading to more congestion and emissions, are of significant concern. The purpose of this document is to provide a high-level framework for assessment of the impacts of road traffic automation.

Members of the Trilateral Working Group on Automation in Road Transportation (ART WG)[1] are working to address the complexity of AV impacts. European researchers are looking at the possibility of applying the Field Operational Test Support Action's framework (FESTA, FOT Net 2016) to automation and sketching the mechanisms through which automation potentially affects our lives. The United States Department of Transportation (US DOT) has sponsored development of a modelling framework that includes the areas of safety, vehicle operations, personal mobility, energy/emissions, network efficiency, travel behaviour, public health, land use, and socio-economic impacts. Japan is developing models of $CO_2$ impacts and has started large scale field operational tests in 2017 under SIP-adus.

To coordinate the impact assessments performed in the field of automated driving, the ART WG established an Impact Assessment sub-group in 2015. The motivation was the realisation that, as field tests are expensive and mostly done on a small scale, international harmonisation would be in everyone's interest. With a harmonised approach, tests and studies can be designed to maximise the insight obtained and to arrange complementary evaluation across the world. Harmonisation would also facilitate meta-analysis.

The framework aims for high-level harmonisation of impact assessment studies globally. It is the first attempt to do harmonisation by the three regions (EC, US and Japan). As there are so many concepts of automated driving, the framework does not give detailed methodological recommendations (i.e., methods to apply for calculating the impact) but it aims to facilitate meta-analysis across different studies. Therefore, the focus is on providing recommendations on how to describe the impact assessment study in a way that the user of the results understands what was evaluated and under which conditions.

The framework presented in this document includes some new material but is partly based on both the US DOT and FESTA frameworks. It also draws from insights obtained at the following workshops:

- FOT-Net Data workshop 'A common methodology for automation FOTs and pilots', Leeds, UK, February 2016
- ITS European Congress, Stakeholder workshop SW3 'Towards a methodology for Field Operational Tests (FOTs) for automated vehicles', Glasgow, UK, June 2016

[1] The European Commission (EC), the United States Department of Transportation (USDOT) and the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) of Japan have a long history of cooperation on Cooperative Intelligent Transportation Systems (C-ITS) activities. The Trilateral Automation in Road Transportation Working Group (ART WG) was established by approval of the Steering Group in October 2012.

- AVS2016 Impact Assessment Break-out session, San Francisco, US, July 2016
- AVS2017 Poster presentation, San Francisco, US, July 2017
- TRB, Poster presentation, Washington DC, US, January 2017
- ITS World Congress, Melbourne, Australia, October 2016
- ITS World Congress, Montreal, Canada, October 2017
- SIP-adus, Tokyo, Japan, November 2016
- SIP-adus, Tokyo, Japan, November 2017
- EU CAD Conference, Brussels, April 2017

There are two major audiences for this report: designers of field operational tests (FOTs) and policy-makers. FOT designers may use it early in the systems engineering process (concept exploration[2], concept-of-operations[3] development, as well as defining the aims objectives, research questions and hypotheses[4]). The framework facilitates starting with the end in mind[5]. For FOT designers, the framework provides a structure for addressing the “where”, “what” and “why” of the project. Section 1.2 of the framework describes the elements of AV system classification including, but not limited to, the operational design domain (the where and what). Later, the elements of the framework itself (section 1.3) help describe “why” the project is being done. The associated key performance indicators (KPIs) provide initial thoughts on measures for validation, to define the data that should be collected, and to ensure that the information gathered maximises the value of the test. Those performing impact assessment for the automation of road transportation can use it as a starting point in design of their evaluation work.

Policy-makers may use the framework to support policy analysis, long-range scenario-based planning, and major infrastructure investment decisions, where various automation futures are envisioned. For policy-makers, the direct and indirect impact areas, as well as their associated linkages, provide a path from the results of a field test, towards potential larger societal impacts. As automation is deployed, the framework may be applied to evaluate the new data that becomes available, and can provide insight as to what related data should be collected.

Finally, for both FOT designers and policy-makers, the framework can support exploratory analysis. For example, users can take broad assumptions about either inputs or outcomes in the future and trace them back through the framework to other things that should be considered or measured. A specific example of the latter might be to consider different roles of shared mobility in relation to transit (ranging from effective last-mile service to full replacement) and mapping that back out to total trips, new types of bottlenecks (e.g. at pickup/drop-off points), and other aspects of demand formation.

This impact assessment framework gives recommendations in Chapter 2 for classifying automation implementations and determining impact areas to be assessed. Chapter 3 presents the impact mechanisms through which automated driving is expected to impact our life, covering both direct and indirect impacts.

[2] Section 4.2.1 of ITS Systems Engineering Guide (National ITS Architecture Team, 2007)
[3] Section 4.3 of ITS Systems Engineering Guide (National ITS Architecture Team, 2007)
[4] Activity 2 in the FESTA framework (Barnard et al., 2017)
[5] See “Start with Your Eye on the Finish Line” in Section 3.2.2 of ITS Systems Engineering Guide (National ITS Architecture Team, 2007)

Chapter 4 provides recommendations for experimental procedure and Chapter 5 for data sharing. Conclusions are made in Chapter 6.

This document is distributed in the interest of information exchange. It does not constitute a standard, specification, or regulation.

## 2 System and Impact Classification

This section describes a framework for assessing the impacts of AV applications. It explains direct and indirect impacts, and the importance of classifying the system before launching into detailed analysis. It then explains each type of impact area in further detail, including providing a list of proposed key performance indicators[6] (KPIs).

AV impacts may be divided into two large groups: direct and indirect. Figure 1 depicts the impact areas. Direct impacts are those which have a relatively clear cause-effect relationship with the primary activity or action. They are generally easier to capture, measure and assess, and are often (though not always) immediate to short-term in nature. In Figure 1, they are in the upper left, and include safety, vehicle operations, energy/emissions and personal mobility. The others are indirect impacts. Indirect impacts can be characterised as secondary, tertiary, or still further removed from the original direct impact. Indirect impacts summarise the broader effects of the individual direct impacts and are produced as the result of a path/chain of impacts, often with complex interactions and external factors. They are typically more difficult to measure and are longer than the time horizon of a field test.[7]

In Figure 1, forward links are represented in green arrows, these impacts are those in which short term changes in an impact area affect longer term changes in another impact area. For example, a change in personal mobility, such as an increase or decrease in shorter walking trips will overtime affect the longer term impacts of public health. Feedback links are represented in brown, these impacts are those in which more holistic or wider reaching changes in impacts affect shorter term and more localised changes in other impact areas. For example, a change in land use and zoning policy will affect the options available for travellers to make personal mobility choices. Finally, those links going in both directions are in black and denote a mutual impact relationship.

[6] Key performance indicator (KPI) is a quantitative or qualitative indicator, 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. (FESTA, FOT-Net 2016)

[7] This explanation is inspired by that of direct and indirect environmental impacts of road development in Roads and the Environment - a Handbook (World Bank 1997)

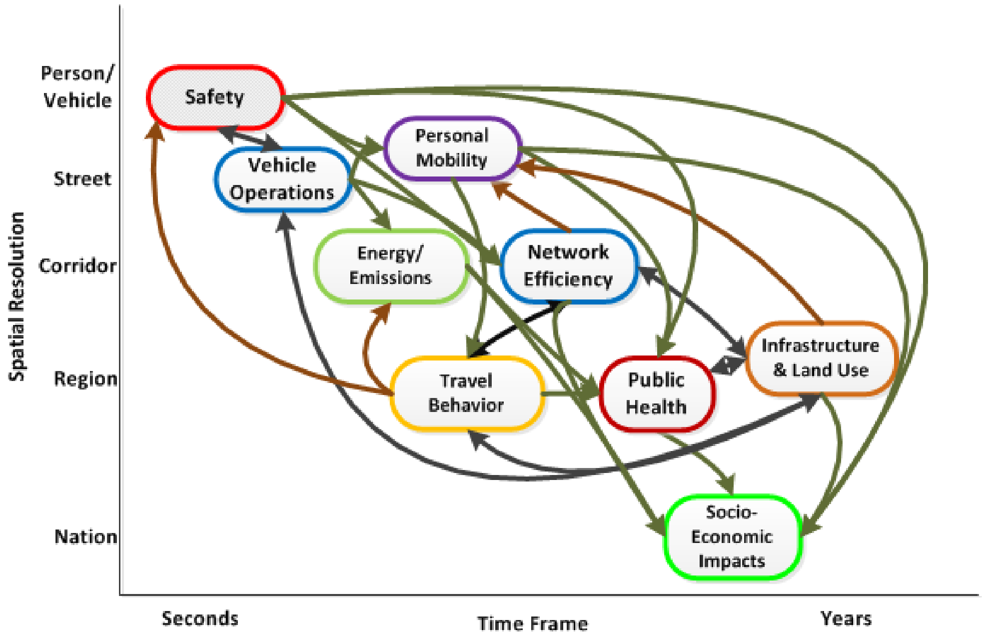


**Figure 1. Impact Areas (Smith, 2016)**

Milakis et al. (2015) discuss first order (i.e. direct) and second and third order (indirect) impacts and linked them to the "ripple effect" used for 'sequentially spreading' events in economics and other social science fields. They note, however, that the parallel is not exact: as shown in Figure 1, there can be feedback from higher-order to lower-order impacts, and as that implies, impact orders may not be in sequence chronologically.

## 2.1 Classification of the system and design domain

As a first step in setting up the analysis of a given implementation of vehicle automation, it is important to specify the system and the service for which impact assessment is made. Otherwise, other researchers risk comparing very different services even though they have a similar design domain. FESTA Handbook (FOT-Net 2016) addresses this classification through specifying use cases. Description of the ART system or service should include (at least):

- Function(s) within the transportation system (e.g., passenger, goods, mixed service, short v. long trips)
- Vehicle type(s) (e.g., private or shared passenger vehicle, mini bus, large bus, truck, etc.)
- SAE level of automation and available automated driving functions
- Operational design domain as defined by SAE (2016)
- Penetration rate of the technology (AVs only or mixed traffic)
- Environment in which the ART system operates (for which the impacts are assessed) (e.g. urban street network, rural road, motorway network, region, etc.)

An example of a classification process from the CityMobil2 project for automated transport systems is shown in Figure 2. A list of elements from this example, specific to public transport, follows. In planned FOTs, a similar classification exercise should be completed for other types of services to ensure impact assessment results are presented in the appropriate context for applying results, harmonisation, and for future meta-studies.

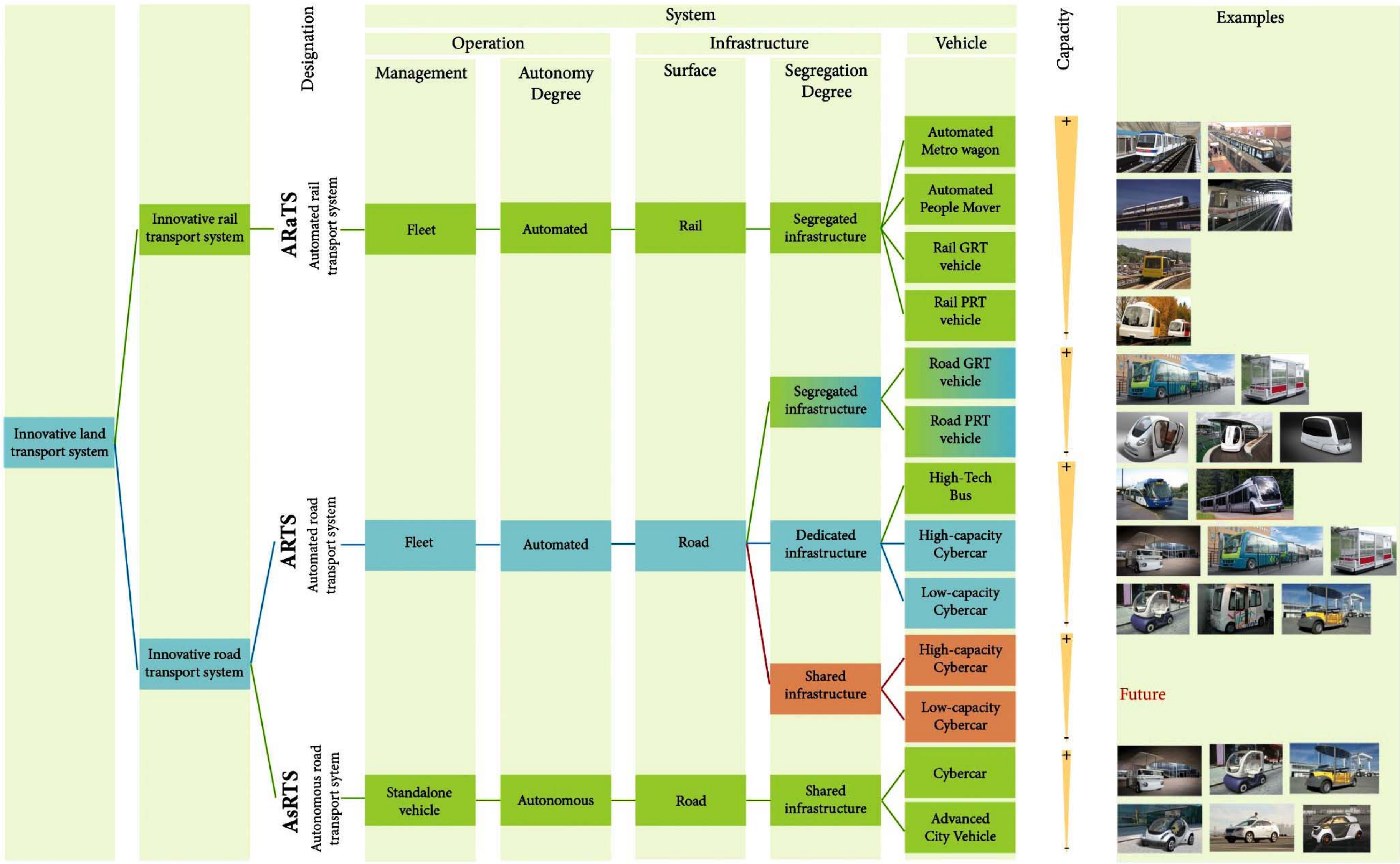


**Figure 2. Automated transport systems classification according to CityMobil2 project (Kyomans et al. 2013).**

Figure 2 classifies the transport systems by:

- Need for mechanical guidance: whether the system relies on guidance from the infrastructure, such as rails. Automation tends to blur the boundaries between road and rail making possible unprecedented synergies (e.g., small automated public transport vehicles could use the metro tunnels at night when it is no longer practical to keep the metro running).
- Level of operation: whether the vehicles are part of a fleet, and whether they need some degree of control from a control system. If the vehicles need external control they are automated but not autonomous; if they can take any decision without any form of communication or cooperation, they are autonomous.
- Degree of segregation of infrastructure: the infrastructure can be either fully segregated (accessible only to vehicles which are part of the system and protected against external intrusion); dedicated (certified to be accessible by the vehicles of the automated transport systems and also to some other users who will need to follow specific rules); or shared (any road infrastructure shared by any road user). Depending on this criterion several different transport systems become possible (e.g. the difference between personal rapid transport (PRT) and the Google autonomous car is mostly there and in fleet supervision). This criterion might even be applied only temporarily giving birth to other new forms of transport like a shared car featuring advanced driver assist systems when in use with a driver on shared infrastructure which could become completely automated for relocation purposes when driving itself on dedicated infrastructures such as bus lanes.
- Vehicle size: this might even be further varied by coupling more vehicles either without mechanical connection (platooning) or with such connection (convoying).

In addition to classifying the system, it is also essential to capture specific information on the operational design domain, including the infrastructure. Under FESTA, these elements are described as situational variables. Elements of the operational design domain include:

- Specific location where the automation system may operate
- Driving mode: examples include expressway merging, high-speed cruising, and low-speed traffic jam
- Level of mapping needed where the automated system operates
- V2X infrastructure
- Type of road: number of lanes and carriageways, required markings, pavement type
- Types of intersections: merge, diverge, traffic signal, stop/yield sign
- Usage of road: exclusive to AVs, shared with other motor vehicles, shared with bicyclists and pedestrians
- Design speed
- Daytime / night time
- Types of road surface conditions: dry, wet, snowy, icy
- Visibility: clear, rain, snow, fog
- Temperature, atmospheric pressure

Classification and description of the system, infrastructure, and operational design domain enables more precise impact assessment and meta-analysis. Recommended direct and indirect impacts for investigation are listed below. In addition, while not impacts per se, other aspects of automation that should be assessed in a field test include the cost and the driver / other road user response to automation. These are also discussed in section 2.2.

## 2.2 Direct impacts and key performance indicators

Direct Impacts are those that can be measured in an FOT. They then can be scaled up from individual field experiment and study levels to a regional or national level, and can lead to indirect impacts. For example, an FOT can measure driving conflicts (Safety), driver/traveller behaviour, car following and intersection performance (Vehicle Operations), energy consumption and tailpipe emissions (Energy / Emissions), and the comfort of the user or the user's ability to multi-task while in the vehicle (Personal Mobility). FOTs can also provide insights into the infrastructure requirements of an automation application.

The Trilateral Impact Assessment Sub-Group conducted an international survey from June-November 2017 where 77 experts representing research organisations, policy makers and authorities, industry and consultants from Europe, the US and Japan provided their views on the importance of different KPIs and proposed additional KPIs to those identified in the survey. This framework uses the results from this survey (Innamaa & Kuisma 2018) to recommend three KPIs for each impact area. Some KPIs are important for more than one impact area, and are listed under each relevant area. Additional KPIs can be found in Annex I.

The following direct impacts are examples of what should be considered in the design of an FOT. For example, specialised use cases such as truck platooning may require a unique set of KPIs not applicable to other FOTs (e.g., number of platoons formed or what share of the vehicle km/miles was driven as part of a platoon).

**User**: In an SAE Level 1-3 system where the driver can choose to use the automated driving functions, the use of automated driving affects the extrapolation of the impacts into a larger context. The use includes aspects like availability, actual use, usability, on which kind of journeys or environments and in which kind of circumstances, and if relevant, what parameters they choose (for example, car following distance). The driver's degree of engagement with the driving task is also relevant (for example, is the driver treating a SAE Level 2 system as though it is a SAE Level 4 system?). For applications operating in mixed traffic environments, the behaviour of other road users (drivers, pedestrians, bicyclists) is also relevant. The following KPIs are recommended to be used:

- Number of instances where the driver must take manual control per 1000 km or miles
- Use of automated driving functions (% of km of maximum possible use)
- Comprehensibility of user interface (expressed on a Likert scale[8], e.g. 1–9, low–high)
- Feeling of safety (expressed on a Likert scale, e.g. 1–9, low–high)

**Vehicle Operations:** Vehicle (control) operations include acceleration, deceleration, lane keeping, car following, lane changing, gap acceptance: all affect road (network) capacity. Relevant automation applications include those which provide longitudinal and/or lateral control with respect to the road and other vehicles. The following KPIs are recommended:

- Number of instances where the driver must take manual control per 1000 km or miles

[8] A Likert scale widely used approach for scaling responses in surveys, e.g., 1 = worst…9 = best.

- Mean and maximum duration of the transfer of control between operator/driver and vehicle (when requested by the vehicle AND when turning automated driving system on/off, i.e. manual overrule)
- Number of emergency decelerations per 1000 km or miles

**Safety:** Ultimately, safety is measured as fatalities, injuries and property damage for vehicle occupants and other road users. Other road users may include pedestrians, bicyclists, slow-moving vehicles, construction workers and first responders. Nearly all AV applications, ranging from SAE Level 1 collision avoidance systems to SAE Level 5 self-driving vehicles, have potential safety impacts. A challenge with safety assessment is that actual crashes are rare events; therefore, proxy measures are often used. These measures may include selected traffic violations, instances where a human driver must take control of the vehicle, exposure to near-crash situations, and responses to near-crash situations. The following KPIs are recommended:

- Number of crashes (distinguishing property damage, and crashes with injuries and fatalities), in total and per 100 million km or miles
- Number of instances where the driver must take manual control per 1000 km or miles
- Number of conflicts encountered where time-to-collision (TTC) is less than a pre-determined threshold per 100 million km or miles

**Energy / Environment:** The energy and emissions category includes both the energy consumption of the vehicle through a driving cycle, and tailpipe emissions of pollutants including greenhouse gases. The direct energy/emissions impacts come from the change in the driving cycle. Changes in vehicle propulsion (e.g., electric vehicles) or an impact on total kilometres/mileage driven may also have a significant effect on tailpipe emissions. The following KPIs are recommended:

- Energy consumption of a vehicle (kWh, litres / 100 km or miles, or electric equivalent[9])
- Tailpipe[10] carbon dioxide (CO2) emissions in total per year and per vehicle-km or mile
- Tailpipe criteria pollutant[11] emissions (NOX, CO, PM10, PM2.5, VOC) in total per year and per vehicle-km or mile

**Personal Mobility:** Mobility from a user's standpoint includes journey quality (comfort, use potential of in-vehicle time), travel time, cost; and whether the travel option is available to someone (e.g., a non-motorist). It also includes equity and accessibility considerations. The higher levels of automation will have the most significant impacts, by providing mobility for non-motorists and enabling multi-tasking. These include first mile / last mile services and accessibility applications. Challenges in measuring personal mobility impacts include the variety of sub-populations who may be affected in different ways, and the difficulty in assessing the actual value of automation to a person based on survey data. (Travel time indicators are generally

[9] 1 gallon (4.54609 litres) of gasoline = 33.7 kWh. See http://www.fueleconomy.gov

[10] In assessing automation benefits, it may be necessary to assume that types of fuel used by automated and non-automated vehicles are the same, at least in the near-term. It may be beyond the scope of an FOT to assess the $CO_2$ emissions from electricity generation.

[11] U.S. criteria pollutants include ozone, particulate matter, carbon monoxide, nitrogen oxides, sulphur dioxide and lead. See http://www.epa.gov/airquality/urbanair/

evaluated at the network level – rather than the individual level. Network efficiency is addressed in Section 2.3 below.) The following KPIs are recommended:

- Type and duration of in-vehicle activities when not operating the vehicle (high levels of automation)
- User perceptions of travelling quality (expressed on a Likert scale, e.g. 1–9, low–high)
- User perceptions of travelling reliability (expressed on a Likert scale, e.g. 1–9, low–high)

**Cost:** Once an automation application has moved out of prototyping, and into production, what is a reasonable estimate of the capital and operating cost for the technology or solution? This is important for assessing the future business case for deployment and ultimate usage. The following KPIs are recommended:

- Capital cost per vehicle and infrastructure for the deployed system
- Cost of purchased automated vehicle (market price, monetary value)
- Operating cost for the deployed system (per vehicle-hour or per vehicle-km or mile, monetary value)
- Investment cost for digital infrastructure (per road km or mile, monetary value)

## 2.3 Indirect impacts and key performance indicators

In assessing indirect impacts, note that together with the digitalisation of transport, there may be changes in service offerings and fleet composition, and these may affect the impacts. For example,

- With better crash avoidance, it may be possible to use lighter-weight vehicles (affects material and energy use or emissions) and avoid congestion (with impact on network efficiency)
- The advanced control systems used for automation may also contribute to electrification (with impact on energy use and emissions)
- If there is no human driver, the layout of the vehicle might change (with impact on energy use and possibilities for non-driving related in-vehicle activities)
- Without the labour cost of a human driver, it may become economical to use smaller vehicles for both trucking and transit (with impact on energy use, network efficiency and possibly for new mobility services).

We are also concerned with how different groups of people might be affected: non-motorists, professional drivers, etc.

**Network Efficiency:** Network efficiency refers to lane, link and intersection capacity and throughput in a regional transport network. It also refers to travel time, delays and travel time reliability. Improved safety may improve network efficiency via reduced incident delay. Also, changes in vehicle operations (e.g., car following) will affect network efficiency. In addition, changes in transport modes or mileage driven by AVs affect it, too. The following KPIs are recommended:

- Throughput i.e. number of vehicles per hour through a particular road section or intersection approach, normalised to number of lanes and proportion of green time (where relevant)
- Maximum road capacity (for a given road section)
- Peak period travel time along a route

**Travel Behaviour:** A traveller may respond to automated transport options, including new service offerings, by changing travel behaviour. There may be more or fewer trips. Modes, routes and destinations may change.

Higher-level automation applications that have a significant effect on personal mobility or labour could have a significant effect on travel behaviour. The following KPIs are recommended:

- Share of transport modes (modal split, based on number of trips)
- Number of trips per week and trip type (in total and per inhabitant)
- Total duration and length in km/miles travelled of trips per week (in total and per inhabitant)

**Asset Management:** Assets include physical and digital infrastructure of road transportation. Automation may affect these infrastructure assets required in several ways, though significant uncertainty still remains in this area. In particular, there may be changes in trip making. If travellers respond to automation by making more trips, more road capacity may be needed.  On the other hand, if automation leads to greater use of shared, rather than owned, vehicles, the infrastructure required for parking may be reduced. Changes in trip making may affect the assets required. Because of this uncertainty, identifying specific indicators is difficult, but the examples listed suggest some areas in which infrastructure assets may be affected. The following KPIs are recommended:

- V2I infrastructure for automation
- Frequency of pothole occurrence (number of potholes per 100 km or miles)
- Use of hard shoulder (for hard-shoulder running or as emergency stop area for mal-functioning automated vehicles)

**Public Health:** Automation may impact the health (physical and mental) of individuals and entire communities, via safety, air pollution, amount of walking and bicycling, as well as access to medical care, food, employment, education and recreation. The following KPIs are recommended:

- Modal share (%) and total mileage travelled (kms) by active modes of transportation (walking and bicycle)
- Number of fatalities and injuries per year per million inhabitants
- Proportion of people with improved access to health services

**Land Use:** Automation may affect the use of land for transport functions (e.g., parking, road geometry) but also in general. These longer-term land use changes may include community planning i.e. location and density of housing, road network design, employment and recreation. The number of factors that contribute to long-term land use changes makes distinguishing those changes contributed by automation a particular challenge. The following KPIs are recommended:

- Number and location of parking slots
- Density of housing
- Location of employment

**Socio-Economic Impacts:** Improved safety, use of time, freight movement, travel options (for motorists and non-motorists), public health, land use and effects of changed emissions (including climate change) will have longer-term economic impacts. Automation may also have substantial impact on labour markets and industries. Assessment in this area continues to evolve. The following KPIs are recommended:

- Work time gained due to ability to multitask while traveling (hours per year, overall and per capita; monetary value)
- Socio-economic benefit-cost ratio
- Work time lost from traffic crashes (hours per year, overall and per capita; monetary value)

# 3 Impact mechanisms

## 3.1 General impact mechanisms for assessment

Nine basic impact mechanisms were formulated for automation studies. The purpose of these mechanisms is to ensure that the assessment covers systematically the intended and unintended, direct and indirect, short-term and long-term impacts of both AV-users and non-users. It is recommended that these mechanisms be used for all impact areas of AD studies.

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 basis for the mechanisms was the nine safety impact mechanisms of intelligent transport systems of Kulmala (2010) which were adapted from the mechanisms formulated by Draskóczy et al. (1998). Kulmala (2010) aimed with his safety assessment framework to eliminate overlaps and thereby the risk of "double counting", to test the validity of any single mechanism, and to operationalise the mechanisms for assessment purposes. The same principles are also valid for automation studies. The aim is to make the mechanisms non-overlapping and all-inclusive, i.e., that all impacts would fall under some and (preferably) only one mechanism. In cases in which an impact falls under two (or more) mechanisms, it is preferable to select the most suitable one.

Kulmala's (2010) safety framework was proposed to be used in ex-ante assessment of in-vehicle safety systems, but also in ex-post evaluation, especially when sufficient accident data is not available to quantify the effects in terms of changes in the numbers of fatalities, injured persons and road crashes. The same goes for the impact mechanisms proposed above. Even if data exists, the systematic way of analysing the cause of impact is always beneficial.

Supporting questions that help to understand what is meant by each mechanism are listed below. Note that all questions might not be relevant for all concepts of AD.

1. Direct modification of the driving task, drive behaviour or travel experience
   - What are the direct impacts due to the vehicle driving by itself (driver?, in-car or remote)?
   - What are the direct impacts of differences in drive behaviour of AV and human driver (in-car or remote) regarding e.g. car-following, target speed, used speed or situation awareness?
   - What are the direct impacts of handover from AV to driver in non-full automation?
   - What are the direct impacts of AV including potentially more active safety systems than the baseline?
   - What are the direct impacts of information and warnings provided by the vehicle sensors to the human driver (SAE 1-2)?

- What are the direct impacts of the non-driving related in-car activities?
- What are the direct impacts on travel experience (e.g. comfort, nausea)?

2. Direct influence by physical and/or digital infrastructure
    - What are the direct impacts due to connectivity (information, warnings, platooning, entertainment, including failures in these: bugs, blind spots, service breaks and hacking)?
    - What are the direct impacts due to having digital maps in use and to the map quality?
    - What are the direct impacts due to physical infrastructure in case it is different for AVs (e.g. special lanes)?

3. Indirect modification of AV user behaviour
    - What are the (long-term) impacts of change in driving skills (in-car or remote)?
    - What are the (long-term) impacts of behavioural adaptation in drive behaviour of the users of AV (when driving in non-AD mode)?
    - What are the impacts of behavioural adaptation when driving in AD mode (long-term impacts on allocation of attention and in-car activities)?
    - What are the impacts of unintended use of AV (e.g. use of AV when not fit to drive; use of low-level AV as high-level)?
    - What are the impacts of failures in connectivity or other features of AV?

4. Indirect modification of non-user behaviour
    - What are the impacts of the behavioural adaptation of the other road users (i.e. imitation of AV drive behaviour)?

5. Modification of interaction between AVs and other road-users
    - What are the impacts on interaction (communication, resolution of encounters) between the AV and other road users (on links and in intersections)? E.g.,
        - Between connected AVs
        - Between connected AV and connected non-AV
        - Between AV and non-connected vehicles
        - Between AV and other road users
    - What are the impacts of the new forms of interaction (including teasing of AV[12]) between the AVs and other road users?

6. Modification of travel behaviour (exposure / amount of travel)
    - What are the impacts (via change in vehicle miles/kilometres travelled or hours travelled) on the number of journeys, e.g., due to impact on comfort and ability to travel?
    - What are the impacts on the length of journeys?

7. Modification of travel behaviour (mode choice)
    - What are the impacts on use of different transport modes / transport mode share (e.g., due to impact on costs, availability, attractiveness, ease-of-use or security of different modes)?

8. Modification of travel behaviour (route choice)
    - What are the impacts of AVs' routes (road type, level of congestion) being different from those of the baseline?

[12] e.g. by jumping to the front of the AV (just for 'fun') to make it stop

9. Modification of consequences due to different vehicle design
   - What are the impacts of the AV design being different from the baseline (outside design, inside design, engine)?
   - What are the impacts of AV including more passive safety systems than baseline?

Interactions with other road users (Mechanism 5) are an essential component of driving activity. In many cases, for example before a lane change or before a left turn with oncoming traffic, drivers often interact with other traffic participants in order to purposefully agree on a future motion plan. Currently, human drivers communicate their intent and anticipate others' intent based on explicit communication means, for example flash of headlights, direction lights, horn, and on implicit cues, for example speed variation, lateral position variation. Similar means are used by pedestrians and other traffic participants to anticipate drivers' intent. Such interactions are expected to gradually change, when the interacting agent is an automated vehicle and not a human driver, following the functionalities and capabilities of automated vehicles.

Table 1 provides some examples of what each mechanism could mean in different areas of ART impact assessment studies.

**Table 1. Examples of different mechanisms per impact area**

| Mechanism | Impact area | | | |
|---|---|---|---|---|
| | **Safety** | **Personal mobility** | **Environment** | **Efficiency** |
| 1. Direct modification of the driving task, drive behaviour or travel experience | Impact on crashes via changes in<br>• situational awareness<br>• perception<br>• speed<br>• car-following behaviour<br>• reaction times<br>due to direct changes in drive behaviour (AV, human driver and in handover) or active safety systems or vehicle sensor based information systems (in non-AD mode) | Impact on travel quality via changes in travel experience, i.e.<br>• comfort<br>• stress<br>• nausea<br>• possibility for non-driving related in-vehicle activities<br>due to direct changes in drive behaviour | Impact on emissions via changes in<br>• speed patterns<br>• acceleration / deceleration<br>• car-following behaviour (especially platooning)<br>due to direct changes in drive behaviour | Impact on throughput due to direct changes in<br>• speed patterns<br>• car-following behaviour<br>• gap acceptance |
| 2. Direct influence by physical and/or digital infrastructure | Impact on crashes via same changes as above due to connectivity and infrastructure quality | Impact on travel quality due to<br>• connectivity (direct impact e.g. on uncertainty and infotainment)<br>• direct changes (e.g. nausea) due to changes in the smoothness of travel due to differences in quality of physical infrastructure | Impact on emissions via same changes as above due to connectivity and differences in infrastructure quality | Impact on throughput due to connectivity, infrastructure quality and change in the use of carriageway width |
| 3. Indirect modification of AV user behaviour | Impact on crashes due to long-term behavioural adaptation of AV user (e.g. reallocation of attention resources, ability to drive / take over driving task) | Long-term impact on amount of travel (number and length of journeys) and used modes of transport via travel quality | Impact via long-term behavioural adaptation of AV user, changes listed above in manual mode | Impact via long-term behavioural adaptation of AV user, changes listed above in manual mode |
| 4. Indirect modification of non-user behaviour | Impact on crashes due to imitation of AV drive behaviour by non-AV users | Long-term impact on used modes due to change in feeling of safety of the non-AV users and due to change in social norms (if everyone else uses, so must I) | Impact on emissions due to imitation of AV drive behaviour by non-AV users | Impact on throughput due to imitation of AV drive behaviour by non-AV users |

| Mechanism | Impact area | | | |
|---|---|---|---|---|
| | **Safety** | **Personal mobility** | **Environment** | **Efficiency** |
| 5. Modification of interaction between AVs and other road-users | Impact on crashes due to change in detection and situation interpretation of other road-users | Impact on travel quality due to smoothness of interaction with other road users (e.g., does the interaction create frustration by failing to meet expectations? ); Impact on amount of travel (travel time) due to changes in the interaction. | Impact on emissions due to change in the smoothness of traffic flows caused by the difference in the interaction with other road-users | Impact on throughput (especially in intersections and on ramps) due to the difference in the interaction with other road-users |
| 6. Modification of exposure / amount of travel | Impact on crashes due to change in distance travelled caused by change in destinations or number of trips due to perceived change in travel comfort and ease | Impact on amount of travel due to change in total time spent travelling due to perceived change in travel comfort and ease | Impact on emissions due to change in distance travelled caused by the change in destinations or number of trips due to perceived change in travel comfort and ease | Impact on throughput due to change in distance travelled caused by the change in destinations or number of trips due to perceived change in travel comfort and ease |
| 7. Modification of mode choice | Impact on crashes due to change in mode selection (risk levels of different transport modes) caused by changes in relative attractiveness and costs of different modes | Impact on used modes due to change in relative satisfaction or attractiveness and costs of different modes | Impact on emissions due to change in mode selection (emission levels of different transport modes) caused by changes in relative attractiveness and costs of different modes | Impact on throughput due to change in modal share on traffic flow (different vehicle types' use of capacity e.g. as passenger car units) |
| 8. Modification of route choice | Impact on crashes via changes in the used road types (their relative risk levels) | Impact on travel quality due to satisfaction on used routes | Impact on emissions due to different speed patterns on different routes (road types) | Impact on throughput via distribution of demand on road network (capacity utilisation rate on different roads) |
| 9. Modification of consequences due to different vehicle design | Impact on crashes via differences in passive safety systems of vehicles eliminating and mitigating crash consequences | Impact on travel quality via differences in vehicle design (e.g. seating, noise) | Impact on emissions due to differences in aerodynamics and engine design | Impact on throughput due to changes in sizes (length) of vehicles |

## 3.2 Impact paths for ART

As there are different levels and concepts of automation, no single approach can be recommended for all impact assessments. Impact paths are introduced to describe a stepwise relation between the operation of the automated driving system, to direct impacts and KPI's. However, the charts below indicate potential impact paths starting from direct impacts on vehicle operations, driver or traveller, quality of travel and transport system. The paths are presented per impact areas for some of them: safety, network efficiency, environment and quality of life / equity / health. These graphs are not all inclusive but they can be used as a starting point for systematically determining the impact paths for systems and impact areas which are under investigation. Naturally, there are strong links between impact areas, e.g., safety impacts affecting efficiency and environment. Thus, assessment of indirect impacts is also recommended.

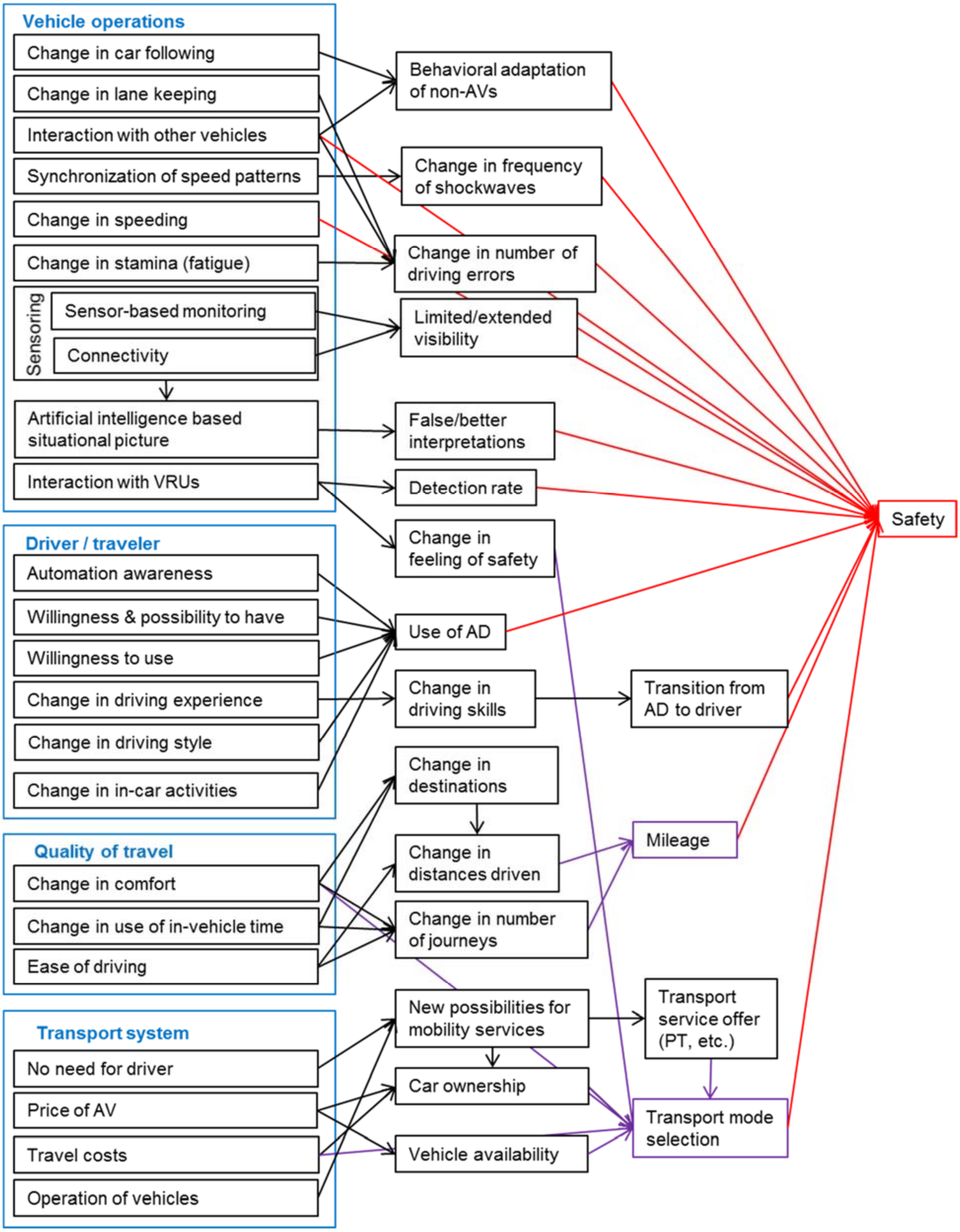


Figure 3. Impact paths of automated driving for safety

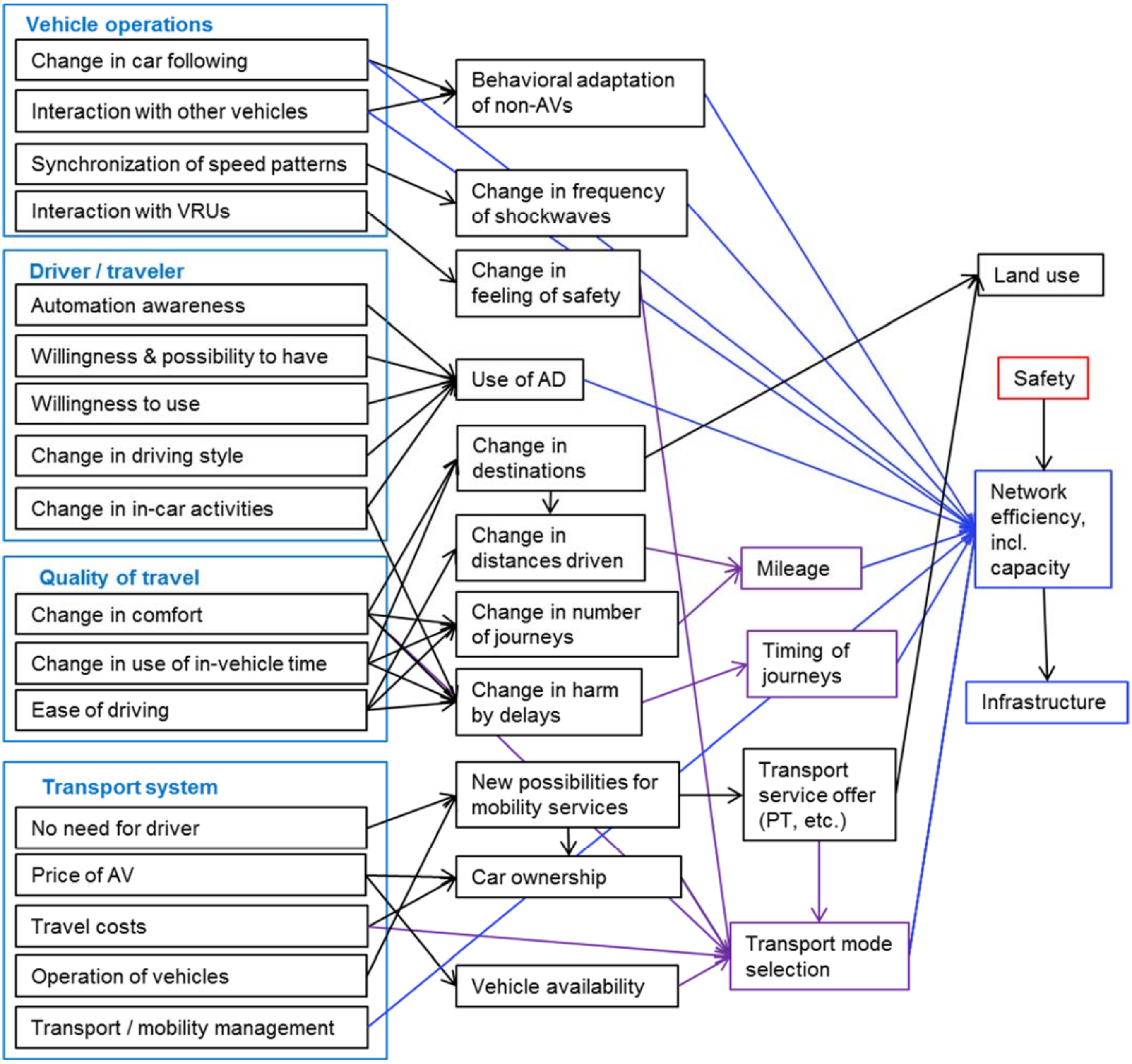


**Figure 4. Impact paths of automated driving for network efficiency including capacity, and infrastructure as well as land use. Note that factors impacting safety are shown in Figure 3.**

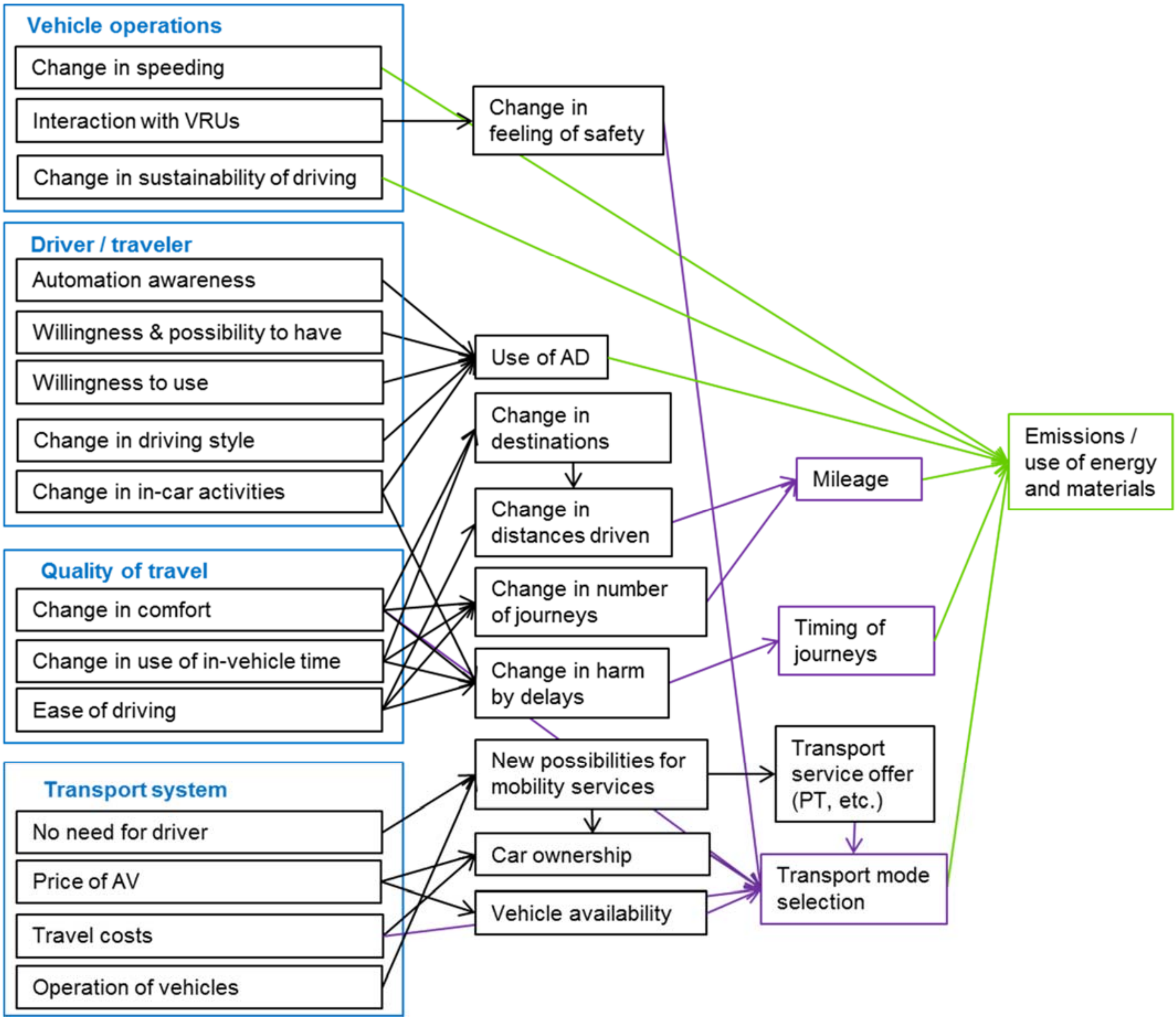


Figure 5. Impact paths of automated driving for emissions and use of energy and materials

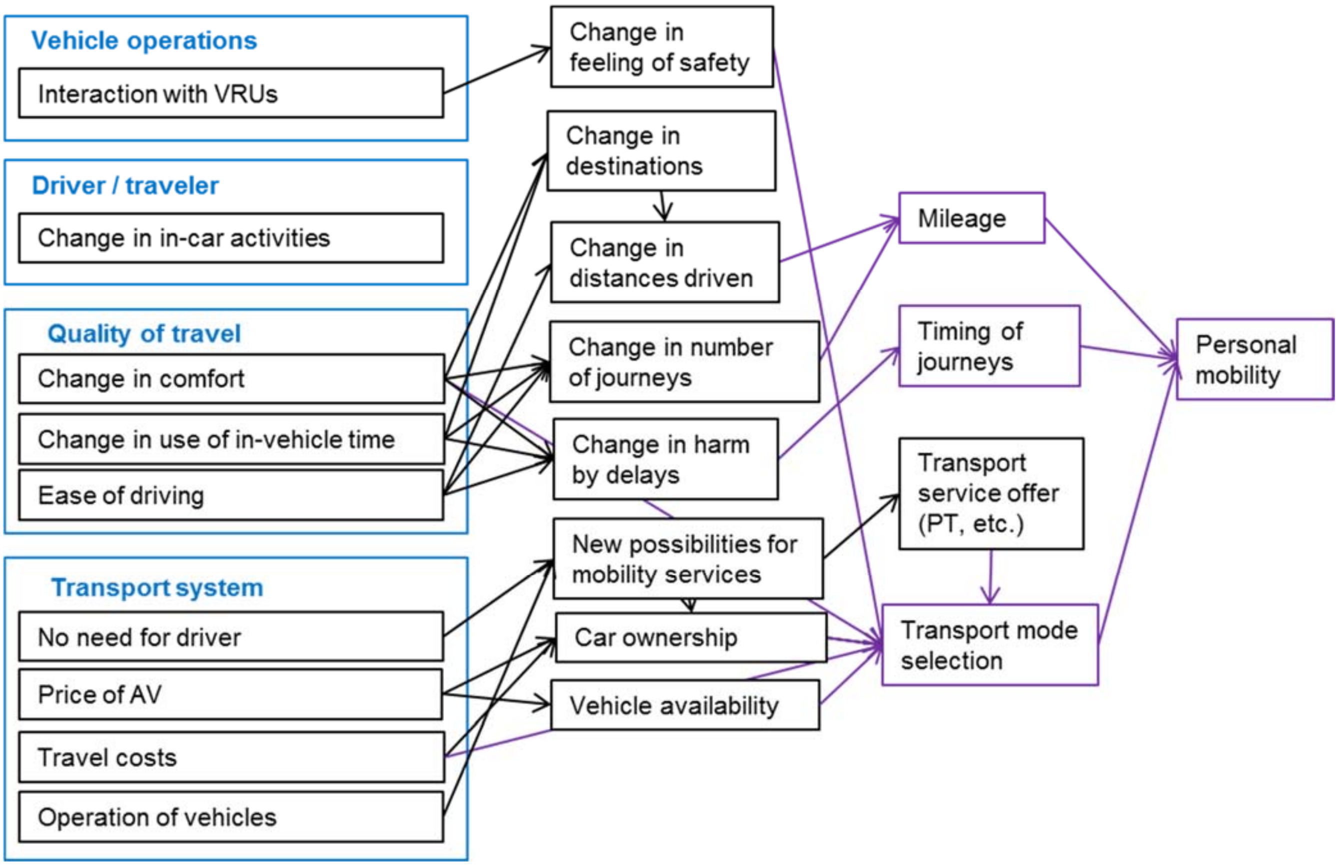


Figure 6. Impact paths of automated driving for personal mobility

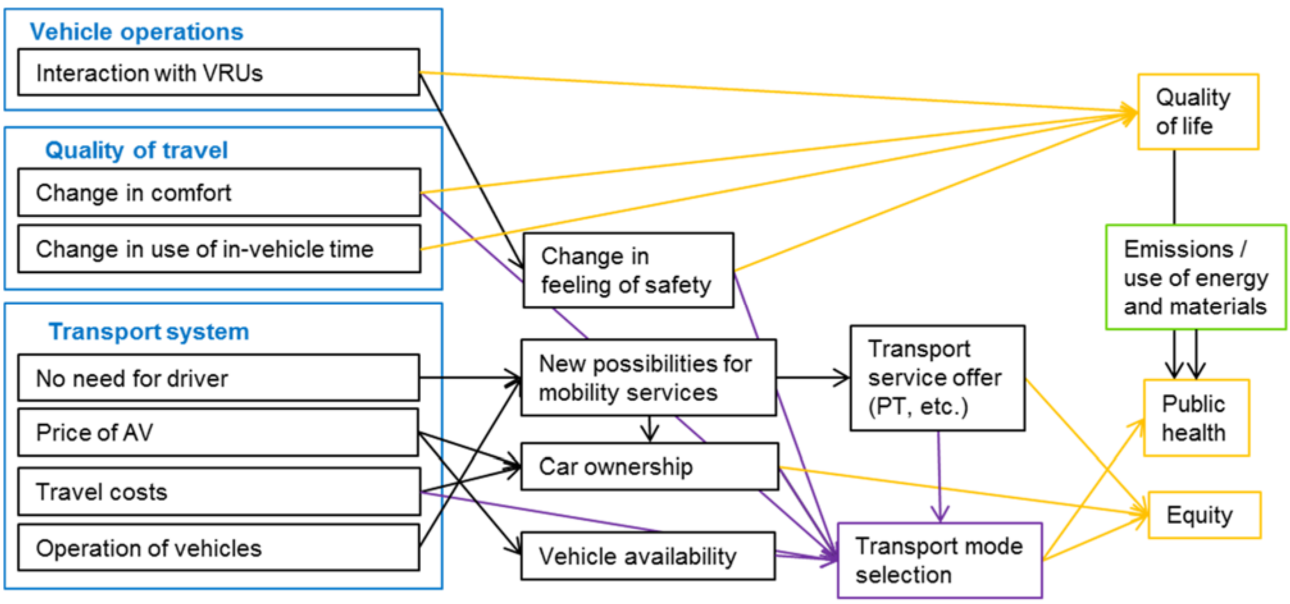


Figure 7. Impact paths of automated driving for quality of life, equity, and health. Note that factors impacting ‘emissions / use of energy and materials’ are shown in Figure 5.

The impact paths presented above should be elaborated further for the system/service under evaluation adding also the direction of change. Figure 8 provides an example for that.

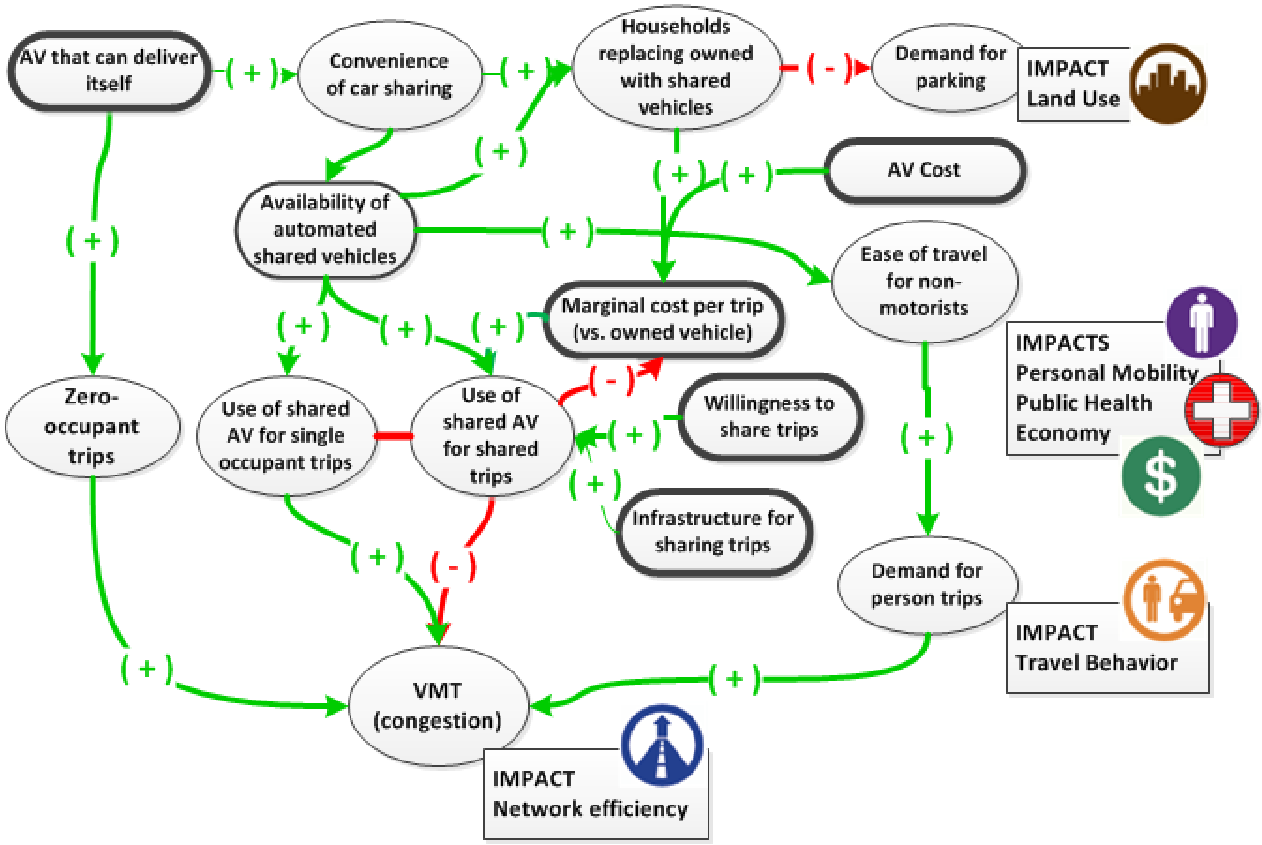


Figure 8. Example of a detailed impact path of 'AV being able to deliver itself'. Note that the green plus signs and red minus signs indicate whether there is an increase or decrease, not whether the change is good or bad.

# 4 Recommendations for experimental procedure

## 4.1 Study design

According to FESTA Handbook (FOT-Net 2016), Field Operational Test (FOT) is "a study undertaken to evaluate a function, or functions, under normal operating conditions in road traffic environments typically encountered by the participants using study design so as to identify real-world effects and benefits". Thus, FOTs are typically used to gather evidence to assess the impacts of automated driving. The FESTA methodology provides an extensive set of recommendations for developing an experimental procedure for FOTs. FESTA is designed for FOTs in general, and does not target AD studies in particular. Yet, many principles from FESTA are relevant here.

The first step when planning an FOT is to identify and specify the concepts (AD functions, systems, or new services) where considerable knowledge about their impacts and effects in realistic (driving) situations is of major interest. For automated driving this step may be a major one, as automation may come in different forms; it could be a suite of automated functions but also completely new types of driverless vehicles or services utilising these vehicles. The description should include technical aspects (like operational area and the control logic); see the description of classification of system and service provided in chapter 2.1). It should also include aspects from the user point of view (such as potential user settings, handover, and human-machine interface (HMI)).

To reach this goal, several steps need to be taken, starting from a description of the automated functions down to an adequate level of detail in order to fully understand objectives and limitations and to derive reasonable use cases. Use cases are a means to describe the boundary conditions under which the automated function is (intended to be) tested; in other words, how we should expect the vehicle to behave under what circumstances. It might also be of interest to define the performance when certain preconditions for the functioning of the AD function (for example when the road markings are not visible) are not met and to identify unintended and unforeseen effects.

Starting from the definition of use cases, specific research questions need to be identified. Research questions are directly related to the impact areas that are to be investigated. First research questions are posed from impact area specific theories point of view (top-down) but also from the system/service under evaluation point of view (bottom-up) to make sure that all relevant questions are considered. Typically, this leads to a list which is too large to be covered by the evaluation. Thus, a prioritisation is needed to select the most relevant research questions which the experimental design can provide the answer for.

The next step is to define statistically testable hypotheses for the prioritised research questions and find measurable indicators to test the hypotheses. Whereas research questions are general questions phrased as real questions ending with a question mark to be answered by compiling and testing related specific hypotheses; hypotheses are statements which can either be true or false. Hypotheses will be tested by statistical means. Defining testable hypotheses may be quite a challenge for road automation studies, as we cannot always predict what the effects are going to be. Hypotheses can only be tested by means of performance indicators. So after establishing these indicators, measures need to be defined and an experimental design to be developed. The following elements need to be determined:

- Characteristics and number of participants
- Characteristics of the experimental environment, such as location, road type, traffic conditions, weather, time of day and season.

It will not always be possible to test automated vehicles in a naturalistic environment. This needs to be addressed in the evaluation. Sometimes a more artificial one needs to be used, such as a location with no other (or limited) motorised traffic or with dedicated lanes. Many AD studies will be performed utilising prototype vehicles, whose performance may not be the same as in the planned production vehicle. There may be failures that will be addressed in the next prototype or production phase.

Most importantly, a lesson learned from all FOTs conducted so far, is that the whole chain, from logging and pilot execution to the analysis, should be piloted before starting the full-scale field tests, allowing for changes in the study design and making sure no technical problems remain.

## 4.2 Baseline

In FESTA Handbook (FOT-Net 2016), it is recommended that driving with the system is compared with driving without it (the baseline). For fully automated vehicles we no longer have a system that can be viewed as independent from the vehicle itself, the whole vehicle is now the system. Some forms of automation, like autonomous or driverless vehicles, mean a radical change in transport, with no baseline available. Comparison with the "old" situation may not be very useful, and studying new emerging patterns may be of more interest. Studying effects against a baseline is important if decisions have to be made on whether the introduction of new systems is desirable. If automation is seen as a process that will continue anyhow, a baseline becomes less important. Similar questions occurred with the introduction of nomadic devices. It is not so interesting to compare the behaviour of travellers without and with a smartphone as their penetration rate went up so quickly that comparison became meaningless, other than for historic reasons. Still, we may want to evaluate the effects of AD in comparison with the current situation. Below we will discuss several options for a baseline.

It might be ideal if we could compare automated driving with non-automated driving (SAE 0). However, there are several reasons why this is not so easy. In the first place, some current vehicles already have some automated functionalities and can be considered to belong to SAE levels 1 or 2. Thus, a traffic flow of 100% non-automated (SAE 0) vehicles is even now only theoretical or belongs to the past. As vehicles and traffic evolve also in other ways than just in relation to automation, comparison to something from the past does not give the true impact compared to the current or future situation.

If the baseline is chosen to be the prevailing vehicle population of the time when the automation FOT is conducted, we will always be comparing different kinds of baseline and a meta-analysis will be difficult to perform. If this option is selected, data and results from Naturalistic Driving Studies might serve as the "baseline." In these studies, "normal" everyday driving is studied (Eenink et al. 2014). This behaviour could be compared with the participants' behaviour when driving in automated vehicles.

If a theoretical traffic flow for a certain level of automation is selected as baseline (any automation level lower than the tested vehicle may be selected), attention must be paid to selecting the automated functionalities of the baseline vehicle or vehicle flow that are meaningful for the impact assessment purposes. A vehicle belonging to certain level of automation may be very different from another vehicle belonging to the same automation level if different aspects of driving have been automated. Thus, if parking behaviour of automated vehicles is studied, the automated parking functionalities are the key criterion in selecting the baseline.

If the comparison is made on a one-to-one level, the baseline vehicle may be a conventional vehicle or a vehicle with lower level automation. However, when impacts are assessed at the level of the entire traffic flow, the penetration of different levels of automation is a key issue, as the penetration of connected or autonomous vehicles as well as the connectivity itself may impact on, for example, the car-following behaviour of vehicles. One could compare situations where there are many automated vehicles in a certain area, versus few or just one. This may require a more experimental set-up, bringing vehicles together, like a study on platooning would do compared to studying single vehicles.

When selecting research questions and hypotheses and building a test design, we must bear in mind that things should not be compared when they are not comparable. For example, the percentage or duration of eyes-off-the-road may be much higher with automation than in conventional vehicles, but is this of any interest? Thus, the baseline should always be selected wisely and guidance from a methodology would be useful.

## 4.3 Controlled testing

Although naturalistic driving studies may be preferred for impact assessment, tests conducted in a controlled environment can also add value.

In case the vehicle with which the tests are made is still a prototype and a so-called safety driver is mandatory to sit beside the test subject of the evaluation, a controlled test approach is practically the only option. The same applies if a large group of test subjects is targeted and the test vehicle fleet is small. If connectivity is evaluated with a small test vehicle fleet, controlled tests are needed to provide a sufficient number of encounters (e.g., with two connected vehicles interacting).

The benefits of controlled tests include effectiveness in data acquisition (i.e. high number of events recorded in relation to the length of data collection period) and smaller variance due to differences in the driving conditions and environments.

The drawback is clearly the short nature of tests per participant. The controlled tests seldom include frequent repetition of AV use for long periods. Therefore, the long-term impacts cannot be seen. This affects e.g. mobility impact assessment as new mobility patterns take time to form. In case a safety driver is on board, his/her presence will also influence the behaviour of the test subject (e.g. you do not start to sleep or mal-use the AD functionalities while being directly observed). Another drawback is the smaller variance in external conditions (weather, road types, traffic situations, etc.).

## 4.4 Use of simulation models

With the indirect impact areas like land use or economy and labour that were not affected or considered in the Advanced Driver Assistance System (ADAS) evaluations and other previous FOTs, new methods need to be developed. In addition to traditional traffic simulation models, econometric and land use models are likely to be relevant.

The use of traffic simulation models can provide the capability to assess the direct impacts of the new/planned ADAS (lower level of automation) onto traffic flows. Since the traffic simulation is often criticised as the black-box, it is required to be modelled with sufficient transparency and rationality supported by the engineering validation process.

In the International Joint Report (IJR) on the "Guidelines for assessing the effects of ITS on $CO_2$ emissions" (Energy-ITS Project, ECOSTAND and PATH, 2013; Figure 9), it is recommended that the simulation models should be clarified with the reference model which figures out the causal relationships of the ADAS functions and the traffic phenomena leading to energy saving and $CO_2$ reduction. The IJR also recommends that the verification process should follow the model clarification to ensure the quantitative relationship of each causal relationship appeared in the reference model, and the validation process to check the model outcomes are similar to fresh real data which might be collected through the small but precise field observations or the coordinated tests.

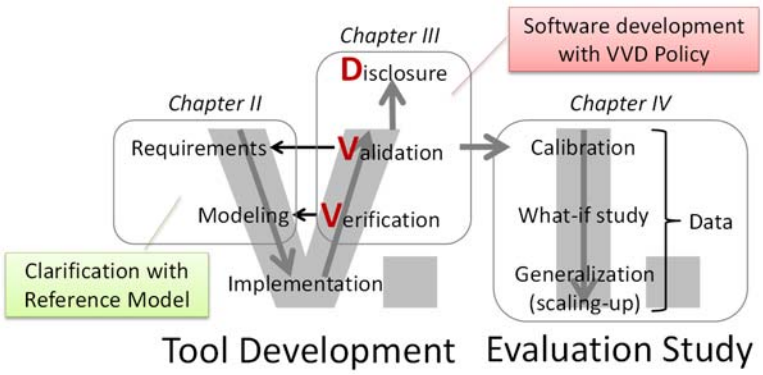


Figure 9. The structure of IJR

There would be the expectation to utilise the simulation models access the social impacts on economics, quality of life, etc. The keys of those assessments are to predict how the ADAS penetrate in the future market and how the travel behaviours of the people will change. The conventional macroscopic modelling using statistics will provide the solution if there is similar and well-studied market under the big assumption that the people will behave as before. Unless we can accept such an assumption, further investigation will be required.

The biggest expectation on the simulation would be to assess the impacts on safety. For this purpose, the microscopic human-vehicle modelling is actively investigated. In general, the modelling tries to identify the mechanism how the human recognition and reaction will lead to the vehicle manoeuvrings and dynamics, and to explain that the ADAS covering the human errors will prevent collisions and accidents. Such direct modelling approaches will provide comprehensive but qualitative answer to the question whether the ADAS will work correctly or not. However, continuous effort is still needed to identify the quantitative relationships in the human model. The use of the driving simulators is encouraged.

Another approach for the safety assessment is to use the microscopic traffic simulation models to estimate the surrogate measures (FHWA, 2001) such as 'time-to-collision', 'post-encroachment-time', etc. By using well-calibrated traffic simulation models in respect of those surrogate measures collected through reliable observations, we may predict how the new/planned ADAS will change the driving behaviours to the safer direction. However, as there is a gap between the changes on those surrogate measures and accident reduction, further investigations with the statistical analysis are still expected. Continuous video survey with effective image processing technology will be the key to collecting massive and precise vehicle motion data.

# 5 Recommendation for data sharing

## 5.1 Reasons for data sharing

It is unlikely that any one study will be able to answer all questions on the impact of road automation nor collect data for that. This means that we will need data from different studies to analyse the wider impact. By sharing data between organisations and projects, new knowledge might be gained about what will happen when automated vehicles drive in real traffic.

From experience in previous FOTs, we know that often large datasets have been collected but because of limited resources, both in terms of time and money, not all data are analysed and impact assessment is often limited. In the US, the Research Data Exchange (https://www.its-rde.net/) has been set up to allow for datasets to be re-used by third parties for new analyses. In Europe, a coordination and support action FOT-Net Data built a Data Sharing Framework (Gellerman et al. 2015) to support data sharing and re-use.

Data sharing and re-using data is a good idea for making efficient use of the large efforts needed to collect the data. It is also allowing researchers who do not have the means to collect new data to answer research questions or providing others with better quality or larger datasets than they could collect. Data from previous FOT and Naturalistic Driving Studies may also be re-used, as this may provide information about the current situation, to be used as a baseline. Also knowledge about how human drivers tackle problems in traffic may provide ideas for informing and improving automated functions.

Sharing and re-using data can be a major step forward in our understanding of the behaviour of transport users and systems. In Europe, the European Commission is stressing this need for data to be open and shared, as they are one of the main sponsors of data collection. A wealth of information is hidden in the datasets that have been collected in recent years and that will be collected on automated driving in the coming years, so the question is how we will be able to generate new knowledge out of these datasets.

Re-using research datasets is seen to have several benefits on a societal level: it will yield further research results at minor additional cost, support education, improve collaboration and thereby create trust in providing more data, and contribute to market introduction of new systems by enabling several organisations to assess benefits. For such reasons, across different fields of science, publicly funded research projects will be required to share more of their collected data in the future.

## 5.2 Obstacles for data sharing and their solutions

Data sharing and re-use is easier said than done. There are five main obstacles:

- Datasets may also contain privacy-sensitive data, such as personal data about the driver, or other data that can be traced back to persons, not only those other road-users participating in the study but random traffic participants happening to be around the vehicles. Specifically, video data are very sensitive. Participants may not have given permission to share data with other organisations and for other purposes.
- Data are not always easy accessible. Only if the dataset is well-documented and contains a rich set of metadata is it possible to create a larger re-use of data collected by others. Not only the data, but also the tools with which the data were processed may be needed.
- Data are valuable and may contain competitive information. Therefore, it cannot always be shared in the format it is collected or at all. Data may contain proprietary information regarding the

performance of the technology used. This area need more attention in automation compared to previous studies due to the responsibility issues within automation.

- Storing, maintaining and opening data after a project has a cost; it may be difficult to find a good financial model for covering or sharing these costs. The owner(s) of the data have also made large efforts (and usually put also their own funding) to collect data and build up the data infrastructure and tools. Data providers and data re-users may have different views on this. It is therefore important to find win-win situations between the data provider and data user in further re-use of the data, to compensate for the efforts made to provide easily accessible data.

## 5.3 Data sharing framework

In the FOT-Net Data coordination and support action, a data sharing framework (Gellerman et al. 2015) has been developed to address these issues. The framework is illustrated in Figure 10.

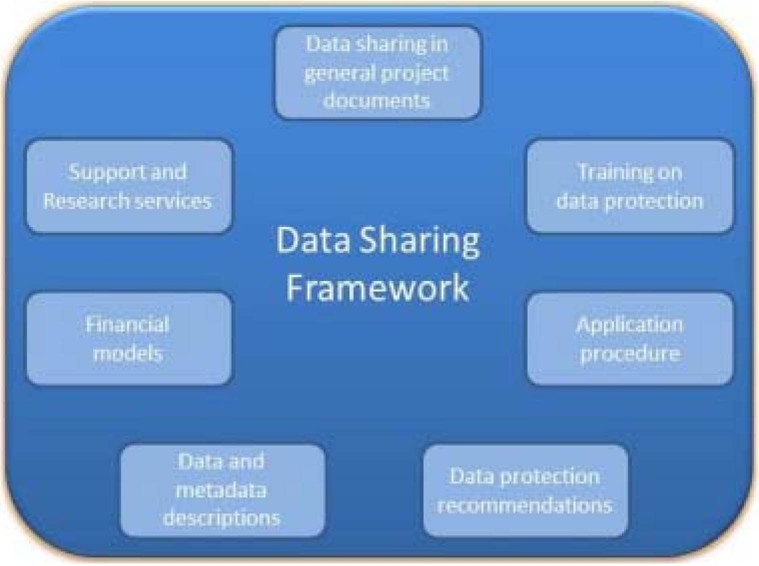


**Figure 10. An overview of FOT-Net's Data Sharing Framework (Gellerman et al. 2015)**

The Data Sharing Framework consists of seven areas, all essential for a smooth data sharing process:

- Project agreements, such as the grant agreement together with the description of the work, the consortium agreement, the participant agreement and external data provider agreements set the pre-requisites and the borders for data sharing together with legal and ethical constraints. The framework provides specific topics to address.

- The availability of documented, valid data and metadata, including a recommendation for a "standard" description of the data, e.g. standard format and the related attributes such as sampling frequency and accuracy, and the description of the study design.

- Data protection requirements both on the data provider and re-users´ analysis site, including security procedures.

- Security and personal integrity training content for all personnel involved.

- Support and research functions, to facilitate the start-up of projects and also e.g. offer processed data for researchers not so familiar with FOT/NDS data. The support also includes the availability of analysis tools.
- Financial models to provide funding for the data to be maintained and available, and access services.
- Last, but not least application procedures including content of application form and data sharing agreement.

## 5.4 Common dataset

The sharing of data from road automation studies and the related analysis would benefit substantially, if a common, minimum dataset could be defined, that studies all over the world collect and share. This does not have to be a large set or include sensitive data. The data probably need to be pre-processed and anonymised to ensure that no confidential or personal information is shared, to facilitate a larger re-use. Examples of data that could be useful to allow cross-study comparison are variables describing vehicle behaviour such as speed, distance to other traffic participants, events and incidents, as well as user questionnaires shedding light on road user acceptance.

It is important to build a common view among the stakeholders of the benefits of such a minimum dataset of automation data. The key is to create a win-win situation, with a picture of the future use of the data resolving questions regarding the impact of automation and removing road blocks for the implementation of automation. The advantages for each of the stakeholders providing and using the data must be clear. The focus should be on sharing data from road automation studies. The advantages probably need to stretch into the deployment phase though, to develop the full picture of potential benefits and create the necessary incentives for sharing data.

The common dataset need to be agreed to be able to be collected and provided by the projects, currently planned and decided.

# 6 Conclusions

This framework is designed to be a high-level evaluation framework for assessing the impact of automation in road transportation and to harmonise these evaluations. The motivation for building this framework is that the potential impacts of automation are far reaching and complex. There are high expectations on the contributions of connected and automated vehicles to societal goals. International harmonisation helps to maximise the insight obtained from single studies, supports meta-analysis and making better use of each other's findings. The framework was developed by members of the Trilateral Impact Assessment Sub-Group for Automation in Road Transportation.

The framework is meant for policy makers to support policy analysis and long-range scenario-based planning, for automakers and after-market equipment manufacturers to better understand the potential benefits of their offerings, for designers of field tests to ensure that the information gathered maximises the value of the test, and for those making the impact assessment as a starting point in design of their evaluation work.

The framework provides recommendation for classifying automation implementations and determining impact areas to be assessed. It presents the impact mechanisms through which automated driving is expected to impact our life, covering both direct and indirect impacts. In addition, it provides recommendations for experimental procedure and data sharing.

This version of the framework is the first full draft to be presented to the target group. Feedback will be collected and later updates will be made to ensure the acceptance and use of the framework in the evaluation studies of road transport automation. Future tasks also include the definition of the minimum common dataset to be collected and shared.

# Acknowledgement

The authors would like to thank the Trilateral Impact Assessment Sub-Group for ART for their contributions to the framework, especially Adriano Alessandrini, Villy Portouli and Ingrid Skogsmo. Risto Kulmala, Anna Schirokoff and Pirkko Rämä are thanked for their contributions when elaborating the impact mechanisms for ART. We would like to thank also the participants of the workshops listed in Chapter 1 for inspiring us to do this work and for giving us new viewpoints. We would also like to thank all those who provided ratings and additional KPIs to the KPI survey.

## References


Draskóczy M, Carsten OMJ, Kulmala R (1998). Road Safety Guidelines. CODE Project, Telematics Application Programme, Deliverable B5.2.

Eenink R, Barnard Y, Baumann M, Augros X, Utesch F (2014). UDRIVE: the European naturalistic driving study. Transport Research Arena 2014, Paris. 10 p. http://www.udrive.eu/index.php/udrive-library/doc_download/24-paper-udrive-the-european-naturalistic-driving-study-in-europe

Energy-ITS Project, ECOSTAND and PATH (2013). International Joint Report on the "Guidelines for assessing the effects of ITS on CO2emissions", http://www.nedo.go.jp/content/100521807.pdf

FHWA (2001). Surrogate Safety Measures from Traffic Simulation Models, Final Report, https://www.fhwa.dot.gov/publications/research/safety/03050/03050.pdf

FOT-Net (2016). FESTA Handbook. Version 6. FOT-Net Data. http://2doubmisw11am9rk1h2g49gq.wpengine.netdna-cdn.com/wp-content/uploads/sites/7/2017/01/FESTA-Handbook-Version-6.pdf

Gellerman H, Svanberg E, Heinig I, Val C, Koskinen S, Innamaa S, Zlocki A (2015). Data Sharing Framework, version 0.5. FOT-Net Data. http://2doubmisw11am9rk1h2g49gq.wpengine.netdna-cdn.com/wp-content/uploads/sites/7/2015/09/FOT-Net-Draft-Data-Sharing-Framework-v0.5.pdf

Innamaa S, Kuisma S (2018). Key performance indicators for assessing impacts of automation in road transportation - Results of the Trilateral key performance indicator survey. VTT Research Report VTT-R-01054-18. 31 + 6 p. http://www.vtt.fi/inf/julkaisut/muut/2018/VTT-R-01054-18.pdf

Koymans A, Limão S, Charlot F, Parent M, Holguin C, Giustiniani G (2013). Functional specifications of vehicles and related services. CityMobil2 Deliverable 15.1. 89 p.

Kulmala R (2010). Ex-ante assessment of the safety effects of intelligent transport systems. Accid. Anal. Prev., doi:10.1016/j.aap.2010.03.001

Milakis D, van Arem B, van Wee, B (2015). Policy and society related implications of automated driving: a review of literature and directions for future research. Delft University of Technology, working paper dated 19 October 2015. Quotation from p. 4.

SAE (2016). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. SAE J3016-2016. SAE International.

Smith S, Bellone J, Bransfield S, Ingles A, Noel G, Reed E, Yanagisawa M (2015), Benefits Estimation Framework for Automated Vehicle Operations, available at https://rosap.ntl.bts.gov/view/dot/4298

Smith S (2016), Benefits Estimation for AV Systems. Presentation at the November 2016 SIP-adus meeting. Available at http://en.sip-adus.jp/evt/workshop2016/file/evt_ws2016_s6_ScottSmith.pdf

World Bank (1997). Road and the Environment. A Handbook. World Bank Technical Paper No. 376. Ed. Tszmokawa K, Hoban C. The International Bank for Reconstruction and Development, THE WORLD BANK. http://siteresources.worldbank.org/INTTRANSPORT/Resources/336291-1107880 869673/covertoc.pdf

(accessed on Jan 4, 2017); http://siteresources.worldbank.org/ INTTRANS PORT/Resources/336291-1107880869673/chap 6.pdf (accessed on Dec 16, 2016).

## Annex I - KPI Repository for impact assessment studies

These lists of KPIs were identified when designing the KPI survey or recommended by those providing answers to it.

**Use of automated driving**

- Use of automated driving functions (% of km of maximum possible use)
- Intended use (statement of interest, % of maximum possible use, by driver, identify relevant journey types)
- Inappropriate use of automated driving functions (number of events per 100 km or miles)
- Reliability (subjective perception, expressed on a Likert scale, e.g. 1–9, low–high)
- Trust (expressed on a Likert scale, e.g. 1–9, low–high)
- Comprehensibility of user interface (expressed on a Likert scale, e.g. 1–9, low–high)
- Requirement of attention and concentration (number of events per 100 km)
- Feeling of pressure because of many parallel tasks (number of events per 100 km)
- Feeling of being able to control the vehicle (expressed on a Likert scale, e.g. 1–9, failure-perfect)
- Mental workload (expressed on a Likert scale, e.g. 1–9, low–high)
- Feeling of frustration (expressed on a Likert scale, e.g. 1–9, low–high)
- Interaction with other road users (expressed on a Likert scale, e.g. 1–9, failure–perfect)
- Feeling of safety (expressed on a Likert scale, e.g. 1–9, very dangerous – very safe)
- Number of instances where the driver must take manual control / 1000 km or miles
- Frequency of miss- or dis-communication with other players, like pedestrians, bicyclists, etc. [per 100km] rate
- Experienced ease of use of an automated vehicle (expressed on a Likert scale, e.g. 1–9, low–high)
- Driver alertness (esp. regarding takeovers related to the SAE Level 3 system)
- Number of instances where the car did not do what the driver expected the car to do (esp. the SAE Level 3 systems)
- Minimum time required to take control (s)
- Number of Incidents per 1000 km or miles where the car must operate in safe mode
- Driver frustration/aggressiveness in the presence of AVs (i.e. frequency of drivers abusing the safety first mechanisms in AVs)

**Vehicle operations**

- Speed variation (st.dev. of speeds) and average speed while travelling at constant speed (on link section, single speed limit)
- Speed distribution
- Proportion of distance driven when speeding
- Lateral position variation (st.dev. of distance from the centre of the lane) while travelling within a lane
- Proportion of correct use of turning indicator/signal
- Mean and maximum longitudinal acceleration and deceleration
- Mean lateral acceleration during lane change
- Number of emergency decelerations per 1000 km or miles
- Maximum jerk (rate of change in acceleration, longitudinal and lateral)
- Mean and minimum distance to the vehicle in front in car following situations (headway 5 s or less)

- Mean and minimum time-headway to the vehicle in front in car following situations
- Variance of the time-headway to the vehicle in front in car following situations
- Minimum accepted gap at intersections or in lane changes
- Mean and maximum duration of the transfer of control between operator/driver and vehicle (turning automated driving system on/off, manual overrule)
- Mean and maximum duration of the transfer of control between operator/driver and vehicle (when requested by the vehicle)
- Number of instances where the driver must take manual control / 1000 km or miles
- Down time frequency (for mechanical servicing/cleaning in the case of ride share autonomous vehicles)
- Number of mechanical/sensor failures per 1000 km
- Number of handovers from autonomous to manual driving at the vehicles' request per 1000 km or miles
- Location of handovers (link/intersection, different road types)
- Frequency of discretionary lane changes (number per 1000 km or miles)
- Number of events per 1000 km or miles when speed needs to be lowered due to other vehicles changing lane in front of the AV
- Frequency of occurrence of TTC (time to collision) below 1 sec

**Safety**

- Number of crashes (distinguishing property damage, and crashes with injuries and fatalities), in total and per 100 million km or miles
- Number of conflicts encountered where time-to-collision (TTC) is less than a pre-determined threshold / 100 million km or miles
- Proportion of time when time-to-collision (TTC) is less than a pre-determined threshold
- Distribution of TTC at brake onsets
- Number of instances with hard braking (high deceleration) / 1000 km or miles
- Number of selected traffic violations / 1000 km or miles of driving
- Number of instances where the driver must take manual control / 1000 km or miles
- Number of false positives / 1000 km or miles, i.e. instances where the vehicle takes unnecessary collision avoidance action
- Number of instances rated by a human as being of increased risk or not correctly handled by the automated vehicle / 1000 km or miles
- Perception of safety by pedestrians, bicyclists, and others sharing the road with AVs.
- Time to take over vehicle control when system cannot provide support / handle the driving situation
- Quality/Type of drivers reaction to a take-over request by the system
- Number of instances when not reacting to a pedestrian appropriately (% of all pedestrians encountered)
- Number of instances when not reacting to a cyclist appropriately (% of all cyclists encountered)

**Energy consumption and environment**

- Energy consumption of a vehicle (kWh/year)
- Energy consumption of a vehicle (litres/100km or miles per gallon or electric equivalent)
- Personal energy consumption (annual average kWh/person-km and kWh/person)

- Total fossil (gasoline, diesel, compressed and liquefied natural gas) energy consumption from highway transportation (tonnes/year)
- Annual traffic CO2 emissions (tonnes/year) on a route or in a region
- Tailpipe carbon dioxide (CO2) emissions in total per year and per vehicle-km or mile
- Tailpipe criteria pollutant emissions (NOX, CO, PM10, PM2.5, VOC) in total per year and per vehicle-km or mile
- Annual average of the proportion of time when noise level above threshold
- Portion of electric vehicles
- Energy use/emissions per second
- BTUs per completed trip -- with trips broken out by length and type of trip.
- BTUs per value of trip (e.g. by trip purpose)

**Personal Mobility**

- Total time spent travelling per day per person
- Number of journeys made per day
- Mean (total) distance travelled per day
- User perceptions of travelling quality (expressed on a Likert scale, e.g. 1–9, low–high)
- User perceptions of travelling reliability (expressed on a Likert scale, e.g. 1–9)
- User perceptions of travelling comfort (expressed on a Likert scale, e.g. 1–9)
- User perception of travel time savings (min per day)
- Type and duration of in-vehicle activities when not operating the vehicle (high levels of automation)
- Types of travellers (children, elderly, disabilities) who are able to use the vehicle without assistance
- Waiting time for vehicle
- Time at interchanges
- Time for detour/load/unload other passengers
- Variation in total time spent travelling and distance travelled per day
- Travel costs (absolute value or as share of personal/household income)
- Reliability
- Trip importance
- Accessibility for disadvantaged or impaired travellers (such as ability to get to essential destinations and ability to get to desirable destinations
- Travel time savings (from distance and time spent traveling)
- Impact on mode choice (e.g. number of journeys made per mode each day)
- User preferences on different modes
- Number of journeys / Distance travelled per mode
- Accessibility to new users (e.g. amount of users for which driving a high level automation vehicle is available)
- Number of new types of trips made per year

**Travel Behaviour**

- Number and type of trips per week (in total and per inhabitant)
- Total duration of trips per week (in total and per inhabitant)
- Total kilometres or miles travelled per week in a region
- Share of transport modes (modal split) per week (based on number of trips)

- Share of used road types per week (based on km or miles travelled)
- Network-level journey time per week
- Average vehicle occupancy rates (persons/veh.)
- Relation of travel times and costs from public Transport, PT-AV-Shuttles and private cars
- Timing of travel
- Vehicle kilometres/miles travelled per person and per vehicle

**Network Efficiency**

- Road capacity at design speed (for a given road section)
- Maximum road capacity (for a given road section)
- Free flow speed (on a given road section)
- Median speed (on a given road section)
- Lowest and highest 5th percentile speed (on a given road section) - addresses "worst case" reliability
- Throughput i.e. number of vehicles per hour through a particular road section or intersection approach, normalised to number of lanes and proportion of green time (where relevant)
- Average travel time (minutes) per road-km or mile
- Peak period travel time along a route
- 95th percentile travel time (minutes) per road-km or mile
- Total travel time and distance travelled per road section or route
- Travel time variability (5th and 95th percentile travel time, to determine certainty in travel time)
- Effective capacity
- Full distributions of travel times and speeds for highway section or network
- Ratios of peak to average travel times and speeds
- Average headway

**Asset management**

- Number of lanes and lane widths
- Use of hard shoulder (for hard-shoulder running or as emergency stop area for mal-functioning automated vehicles)
- V2I infrastructure for automation
- Size and weight implications of changed fleet composition
- Minimum bearing capacity on a road section (tonnes)
- Pavement damage (level of damage, damaged area, % of road km or miles)
- Mean rut depth (mm or inch with 2 decimals)
- Frequency of pothole occurrence (number of potholes per 100 km or miles)

**Costs**

- Capital cost per vehicle for the deployed system (infrastructure, monetary value)
- Cost of purchased automated vehicle (market price, monetary value)
- Average annual maintenance costs of automated vehicles (currency/veh./year)
- Operating cost for the deployed system (per vehicle-hour or per vehicle-km or mile, monetary value)
- Cost per trip (for user, monetary value)
- Investment cost for physical infrastructure (per road km or mile, monetary value)

- Operation and maintenance cost for physical infrastructure (per road km or mile, monetary value)
- Investment cost for digital infrastructure (per road km or mile, monetary value)
- Operation and maintenance cost for digital infrastructure (per road km or mile, monetary value)
- Investment cost for connectivity network (per road km or mile, monetary value)
- Operation and maintenance cost for connectivity network (per road km or mile, monetary value)
- Cost of education per driver (monetary value)
- Cost for retro-fit kits
- Total cost per mile (purchase, maintenance, operation)

**Public Health**

- Quality-adjusted life years
- Population exposure to air pollution
- Number of fatalities and injuries per year per million inhabitants
- Modal share (%) and total mileage travelled (kms) by active modes of transportation (walking and bicycle)
- Proportion of people with improved access to health services
- Proportion of people with improved access to recreation and other services

**Land Use**

- Density of housing
- Road network design
- Location of employment
- Location of recreation
- Number of parking slots
- Location of parking
- Density of employment and shopping
- Creation of new real estate developments or new towns with transportation infrastructure designed specifically for AV access
- Distance in time to employment
- Space needed for road
- Space needed for transport and parking

**Economic Impact**

- Gross Domestic Product (hours per year, overall and per capita; monetary value)
- Total factor productivity / multi-factor productivity estimates
- Work time lost from traffic crashes (hours per year, overall and per capita; monetary value)
- Work time lost from illnesses related to air pollution [hours per year, overall and per capita; monetary value]
- Work time gained due to ability to multitask while traveling (hours per year, overall and per capita; monetary value)
- Labour force participation rate – overall and for non-drivers
- New established businesses / job creation
- Number of vanished/disappeared jobs
- Socio-economic benefit-cost ratio

- Number of Providers of AV-fleets in a local market
- Market share of trips in shared fleet and privately owned cars
- Portion of mobility expenditures of house-hold income