Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 72 (2023) 2054–2061

Transportation Research Procedia

www.elsevier.com/locate/procedia

Transport Research Arena (TRA) Conference

Methodology for Field Operational Tests: updating the FESTA methodology for connected and automated driving pilots

Yvonne Barnard[a1], Sami Koskinen[b]

[a]*Institute for Transport Studies, University of Leeds, Leeds, United Kingdom*
[b] *VTT Technical Research Centre of Finland Ltd., Finland*

Abstract

The FESTA methodology for Field Operational Tests (FOTs) was developed in 2008 to support the evaluation of advanced driver-assistance systems. It has since been updated many times, and successfully used in European and national projects. The paper describes how the methodology was extended in order to be useful for Connected, Cooperative, and Automated Mobility (CCAM) projects. The FESTA handbook was updated with more recommendations addressing automation. A micro-version of FESTA was developed for small-scale pilots. A gap analysis was performed to identify new elements for CCAM evaluation. New approaches for impact assessment were developed. Recommendations for gathering lessons learned and for sharing data, knowledge and experience have been included in FESTA. A new EU project, FAME, will take the next step in evaluating CCAM, developing a Common Evaluation Methodology.

Keywords: FESTA; Connected, Cooperative, and Automated Mobility; Field Operational Tests; Evaluation methodology; Impact Assessment

1. Introduction

Numerous projects have recently set out to test Connected, Cooperative, and Automated Mobility (CCAM) technologies and services. Automated driving tests are gradually branching out from test tracks to a wider and more realistic environment. Projects range from small-scale pilots to large European Field Operational Tests (FOTs). To be able to evaluate these complex systems in a systematic and rigorous way, a comprehensive evaluation methodology is needed. Over the past decade, experience has been gained on the evaluation of advanced driver-assistance systems and other vehicle information and communications technologies. In order to support this evaluation, the FESTA methodology for FOTs was developed in 2008 in the European FESTA project, and described in the FESTA Handbook (FESTA, 2021). It has since received many updates. This methodology has been successfully used in large-scale FOTs, in European and national projects. Projects that evaluate CCAM, however, usually conduct field tests on a much smaller scale, due to the complexity and the novelty of automated systems. However, the principles on which

[1]* Corresponding author
E-mail address: y.barnard@leeds.ac.uk

10.1016/j.trpro.2023.11.688

FESTA methodology is built as well as its practical recommendations can support various types of user and vehicle related tests. The paper will describe how the methodology has been updated and extended in order to provide practical guidance for CCAM projects.

Nomenclature	
CCAM	Connected, Cooperative, and Automated Mobility
CEM	Common Evaluation Methodology
FOT	Field Operational Test
SD	System Dynamics

2. The FESTA methodology

The FESTA methodology as described in the FESTA Handbook draws from experience and knowledge across numerous past Field Operational Tests (FOTs) of vehicle information and communication technology systems. In the Handbook an FOT is defined as:

"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."

The first version of the handbook was created by the European support action FESTA (Field opErational teSt supporT Action, 2007–2008). The handbook has since been frequently updated by the FOT-Net community and follow-up coordination and support actions CARTRE and ARCADE (CARTRE; ARCADE). The handbook gives practical advice on how to set-up and conduct FOTs and how to analyse the results. It provides and advocates a systematic research-oriented approach. FESTA offers a framework with definitions, categories and worked-out examples. It provides helpful checklists for all phases of a project and discusses best practices from past projects. The FESTA methodology is summarised in the so-called FESTA V in Figure 1.

The FESTA V addresses three phases in the evaluation, and a number of steps in each phase. Although this may give the impression that the steps are to be performed sequentially, in reality iteration will be required and is recommended in the Handbook. The three phases are:

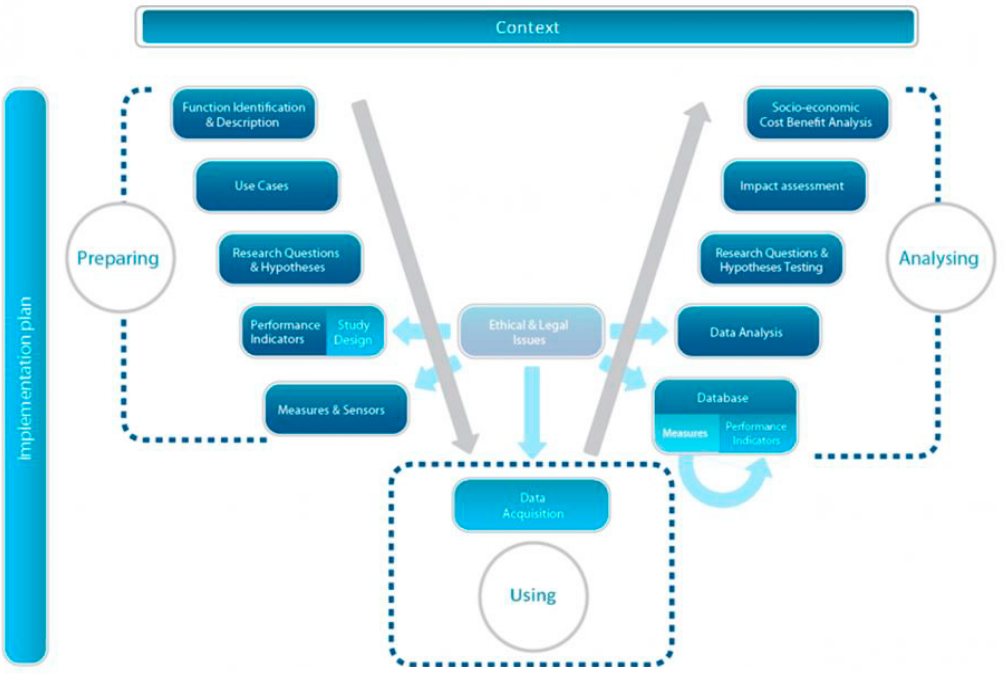

Fig. 1. The FESTA V (FESTA, 2021)

PREPARING: The steps on the left side of the FESTA V highlight that research should start from understanding the tested systems and defining the main research questions. Commonly agreed priority research questions and test plans and methods are key for an efficient collaboration between partners. It pays off to clarify goals and get the team committed. Research questions lead to related data needs and performance indicators, and how to collect such evidence.

USING: Managing vehicle fleets and recruiting test subjects in order to collect data is a feat in itself. Data should be free of major errors and well documented to support analysis by several persons. In addition, one has to ensure the safety and privacy of test subjects and maintain confidentiality regarding the trade secrets of the tested products. Before the tests can truly begin, a phase of pre-testing (piloting) is required to make sure that everything will run correctly and smoothly.

ANALYSING: Measurement data, observations and interviews are used to conclude on the research questions and hypotheses. Effects identified in the tests are considered also from wider societal/stakeholder perspectives. The findings are to be scaled up: what would change, if the tested new system or service gains popularity. Finally, a cost–benefit analysis can clarify the monetary value of identified benefits versus costs.

An important aspect of the FESTA V are the legal and ethical issues that arise at many stages. The context is mentioned at the top of the V: For instance, the choice of a function to be tested implies that there is either a problem that is to be addressed and that the chosen function is defined to solve the problem, or that a policy objective is stated and that the function being tested can be used to reach the objective. The context in which an FOT is being performed is very important, and influences the choices that have to be made in the evaluation and the questions that the FOT need to answer.

3. The update of the FESTA methodology

The ARCADE coordination and support action, funded by the European Commission's Horizon 2020 programme, was responsible for the updating and disseminating of the FESTA methodology (ARCADE). ARCADE aimed at building consensus across stakeholders from all sectors for a sound and harmonised deployment of CCAM in Europe and beyond. This network exchanged knowledge and experiences, built up synergies and a common approach to development, testing, and validation of CCAM.

Evaluating automated systems and services, which are far more complex than single driver assistance systems, requires an adapted methodology. In order to adapt FESTA for the evaluation of CCAM, several steps were taken. The handbook was updated with more recommendations addressing automation. A micro-version of FESTA was developed for small-scale tests, for example national projects. A gap analysis was performed to identify new elements for CCAM evaluation. New approaches for impact assessment were developed.

3.1. Update of the handbook

The FESTA Handbook was updated with recommendations for CCAM tests. First CCAM-related updates of the FESTA Handbook were compiled by the CARTRE project. These updates were discussed in version 7 of the handbook and further ideas at the time have been documented e.g. in Barnard et al., 2016, and in Innamaa et al., 2018. The latest version 8 of the handbook, from September 2021, was edited by the authors of this paper with support from ARCADE project partners. It now includes further guidance and clarifications on a number of aspects: lessons learned from the first CCAM tests are discussed and special attention is paid to the latest developments on legal and ethical issues (FESTA, 2021). The handbook provides guidance on what to consider in an FOT, explains what are GDPR (General Data Protection Regulation) research exemptions, and offers checklists related to drafting consent forms and organising introductory sessions for participants. A new final chapter was added to the handbook, discussing the collection of lessons learned, and sharing knowledge and data.

The term connected and automated vehicles may refer to several levels of automation, ranging from vehicles with systems that may take over some tasks from the driver to completely driverless vehicles. The FESTA methodology is strongly centred on the drivers of vehicles, and the changes in their behaviour when driving a vehicle that is instrumented with new systems. This focus is changing with CCAM FOTs towards the behaviour of the vehicle and the wider mobility behaviour of the user of these vehicles. Automated driving is often closely related to connectivity between vehicles, and between vehicles and the infrastructure. In the handbook, sections are devoted to cooperative systems, as several large FOTs have been performed on the operation of these systems. However, as connectivity

becomes a major feature in automation, not only the behaviour of the vehicles needs to be studied but also the functioning of other connected components and their communications.

An example of issues arising in the methodology when evaluating CCAM, is the formulation of research questions. Experience from FOTs showed that defining and especially selecting research questions is already one of the hardest parts of setting up an FOT. As CCAM FOTs are dealing with new types of vehicle and mobility, as well as with new impact areas, many new questions will arise. It will not be possible to answer all these questions, so choices have to be made.

We may distinguish three types of research questions for CCAM FOTs:

- User-centred tests addressing questions about how the user/driver reacts to automation and uses it. Examples are: how they understand the capabilities of the vehicle; what they will do when not supervising the surrounding vehicles/traffic; the transition from automated driving back to driver driving; the interaction with other road users including vulnerable road users and drivers of conventional vehicles; user acceptance; changes in mobility patterns; etc.
- Vehicle-centred tests addressing the question of how the automated vehicle behaves in different conditions and environments. Here, for example, questions arise about the situational picture that the automated vehicle creates, and the interaction of the automated vehicle with other automated and conventional vehicles, other road users and with the infrastructure.
- Context-centred tests addressing questions of wider impacts. For example: how mobility of different user groups and the services they use change, what the impacts are on the traffic flow level or on the transport system level, how the position of other road-users like pedestrians and cyclists is affected, what ethical choices might be involved, and what would be the impacts on the built-up environment, industry and society. These types of questions are extremely important but not easy to investigate with FOTs as these impacts typically take a longer period to evolve than the duration of an FOT.

Finally it should be noted that field operational testing is not the only way to evaluate automation; it is a method that is complementary to (or is preceded by) simulator and laboratory studies, expert assessments, technical tests on test-tracks, public acceptance studies, etc.

3.2. Micro-FESTA for small-scale tests

Although FOTs have been carried out in large European projects, lasting several years and involving many partners, smaller tests are also taking place. Not all of the detailed steps in the FESTA methodology may be appropriate for these small projects. For small-scale efforts, predominantly run by a couple of organisations and taking place in a single city, Micro-FESTA (Barnard et al., 2021) was developed in the CARTRE and ARCADE projects (CARTRE; ARCADE). Small pilot projects would not have the scope or resources to make full use of the FESTA approaches. They carry out tests with limited means, but do not want to come up with limited results or vague lessons. These projects can offer significant contributions to gathering knowledge on the effects of road automation.

Pilot projects and FOTs actually share the main methodological elements: defining the test(s), setting up data collection, recruiting users, analysis of the data, and considering the results with stakeholders. Use of FESTA lessons learned can boost pilot projects in many phases. In particular, reusing existing tools, templates and learning from best practices can significantly speed up test preparations. FESTA offers a scientific framework with its methods, categories and definitions. It provides practical guidelines for each step of testing, and helps to define a clear setup, aiming to achieve valid and easy-to-understand results. Micro-FESTA introduces the key points of FESTA methodology, and discusses how they can benefit small projects. At the same time, Micro-FESTA could be used as an entry for people who are newly involved in evaluating automated systems.

Examples of recommendations provided are:

- Gather information on the systems and functions to be piloted also from evaluation and user testing perspectives, e.g. which accident types and operations they might impact. Be clear about the technology matureness and practical readiness.
- Find a good balance between technology development and identifying societal/industrial benefits.
- Consider the role of infrastructure and links to other systems: connected and automated systems often require support from other systems or some changes to digital or physical infrastructure to reach their true potential, e.g. higher speed and efficiency.
- Documentation in all stages is key to obtain results that are useful for the future.

- Multi-stakeholder representation is usually needed. Develop a communication strategy with stakeholders and the general public.
- When working with private partners, some data, results and methods can be shared, but not all. Pushing for collaboration often improves the project results, even if non-disclosure agreements are needed, but beware that these may take some time to establish.

3.3. Gap analysis

The ARCADE project has been carrying out a gap analysis to identify the topics and guidelines that are not yet well addressed by FESTA or that arise from the new developments in CCAM evaluation. A (on-line) workshop with some 70 participants was held in November 2020 to discuss methodological gaps and directions for the development of a common evaluation methodology specifically for CCAM (Kalisvaart et al., 2021; ARCADE, 2021).

Examples of gaps related to the three FESTA phases are:

Study design

- Approaches for achieving a realistic and rich user experience with early prototype vehicles
- Method to define and measure a clear baseline for CCAM impact assessment
- Guidelines for study designs enabling generalization of the results (needed for the impact assessment), taking into account the scale of the test

Data acquisition

- Guidelines for efficient and effective process obtaining a test permit for public roads
- Common solutions for data management (release, flow, models, formats)
- Agreed principles for sharing data between industry and research, taking into account the sensitivity of the data from a business perspective

Impact assessment

- Common performance indicators for assessing the impacts
- Shared assumptions about changes in human behaviour as the proportion of CCAM increases
- Shared future scenarios for generalization of impact assessment

Some of the gaps identified will be difficult to address, as evaluation of CCAM requires innovative methods and tools. As CCAM services and technologies may evolve rapidly, there is a need for more agile evaluation processes, without sacrificing the structured and evidence based approach advocated by FESTA.

3.4. Impact assessment

Assessing the impacts is one of the main steps in the methodology. One of the main objectives to perform an FOT is to acquire solid information on the potential impacts a widespread adoption of the evaluated CCAM technologies and services will have. Impacts will be found in many different areas, some close to the traffic system, for example on congestion and road safety, others on wider societal impacts such as energy use, land use and equity (Innamaa et al., 2018; Aittoniemi et al., 2020). In ARCADE a variety of societal challenges and potential scenarios were described (Giro et al., 2019). Thematic areas identified are: policy and regulatory needs, socio-economic assessment and sustainability, safety validation and roadworthiness testing, and user awareness, societal acceptance and ethics.

The European CCAM Partnership under the Horizon Europe programme, aiming to harmonise European R&I efforts to accelerate the implementation of innovative CCAM technologies and services, sees as one of the major problems that the long-term implications, benefits and impacts of integrating CCAM solutions into the mobility system are not sufficiently examined (CCAM partnership). However, impact assessment of such complex mobility developments is not an easy task. FESTA provides recommendations for impact assessment for areas like safety, mobility, efficiency and environment; assessing impacts of CCAM on the wider socio-economic area may require the development of advanced methods.

Impact assessment is also an important challenge in an international context. The Trilateral Working Group on Automation in Road Transportation established a working group on Impact Assessment. The group is an initiative of the European Commission, the United States Department of Transportation and the Japanese Ministry of Transportation. In 2015, the Impact Assessment group began working together to develop a high-level framework for assessing the impacts of road traffic automation. The goal was to facilitate impact assessment work in projects involving field tests with automated vehicles in the three regions and beyond. The framework (Innamaa et al., 2018)

provides recommendations on how to describe impact assessment studies in such a way that the user of the results understands what was evaluated and under which conditions. In addition, a detailed series of key performance indicators was developed. Currently the group works on using System Dynamics (SD) modelling to get a better insight in the impact mechanisms affecting society. System dynamics is an approach to modelling complex problems over a specific time period, capturing time delays, non-linearities and feedback loops (CAD Knowledge Base Methodology; Harrison et al., 2021). The group also works with collective modelling, building consensus models, as a way to engage stakeholders and to elicit their ideas (Rakoff et al., 2020).

4. The role of sharing and learning

A crucial role in keeping the methodology alive and adapting it to new developments is to learn from experiences and share lessons learned. The knowledge base for CCAM, maintained by the ARACDE consortium, contains a large number of lessons learned from European and national projects, many of them related to the evaluation process (CAD Knowledge Base).

The latest update of the handbook added a new chapter about collecting lessons learned. Earlier versions already paid attention to the need to share data, knowledge and experience. As the research community begins to evaluate automated mobility, sharing becomes even more important, as little is yet known about how new technologies and services work out in real life and on public roads, while evidence is needed about the safety and efficiency of CCAM.

Lessons learned may concern many aspects of FOTs, and each step in the FESTA methodology may lead to lessons learned. FESTA strongly recommends to document all steps of an FOT, as this is important for both interpreting the results and learning from the (good and bad) experiences.

We may distinguish several types of lessons learned:

1. *Set-up and execution of a field test*

In setting up the test, many choices and decisions have to be made, for example, selection of research questions and experimental environment. Practical issues have to be solved, such as the availability of vehicles and participants, use of resources, and permissions needed from authorities and ethical committees. Not all problems can be foreseen and new issues may arise during the execution of the field test. Learning from the experiences of other projects may be very valuable in avoiding and mitigating problems.

2. *The systems and functions under evaluation, their description*

A large variety of systems and functions may be evaluated, and a wealth of experience is available from projects, on how to deal with the detailed technical aspects. With fast-paced development, lessons learned may be too specific and will soon become outdated. However, it may be very useful to search for experiences from other projects evaluating systems with similar features.

3. *Data gathered in the project and external data, data processing and storage*

For both data gathered by the project and data from external sources there may be many issues, e.g. with content and quality of the data. Some of them are not known before the analysis process. Increasing the common knowledge on data sources with sufficient quality or at least the weaknesses of different (especially public) data would save time and effort.

4. *The evaluation process*

Lessons learned concern how the evaluation was set up in an FOT, starting from setting goals to reaching the final conclusions. Examples are the way in which certain issues in the project were managed and structured, goals were set and realised, resources managed, and results were interpreted and justified.

5. *Evaluation tools and sensors*

New tools and sensors are rapidly becoming available, so lessons about specific tools may become obsolete at some point. However, it may be useful to search for experiences from other projects before acquiring hard- and software tools and sensors or developing them by the project itself.

6. *The experiences and acceptance of the users*

Performing user tests with complex systems such as automated vehicles in a natural environment may be complex and challenging, and especially so, if the systems are not yet mature and are still under development. Users may even not be allowed to operate the systems themselves, or systems can only be tested in artificial circumstances. Sharing best practices on how to enhance the user experience may be very helpful.

7. *Communication with stakeholders*

The goal of an FOT is usually to produce results that are useful for stakeholders. Lessons learned about stakeholder analysis, managing the expectations of stakeholders, communicating with stakeholders and translating technical results into information that can be used by different groups of stakeholders, are very useful. These lessons are not always easy to share for confidentiality and commercial reasons.

8. *Sharing methods, tools, data and knowledge*

When working with external partners, some data, results and methods can be shared, but not all. Even within a project, some data will remain confidential. Practical solutions found for data sharing within the project and outside the project are useful for others. FESTA encourages FOTs to share their methods, tools, knowledge, and data. Re-use of existing methods and tools can make a big difference in the time that is needed to prepare an FOT. Sharing knowledge and data can offer interesting collaboration and dissemination options (FOT-Net Data and CARTRE EU support actions, 2019).

Finally, lessons can be learned about the methodology itself, both for the purpose of improving the application of the methodology and for improving the methodology.

5. Conclusions and future work

A common methodology adopted by projects has major advantages, as it allows comparison of results between FOTs, provides a common vocabulary, and improves communication between stakeholders involved in FOT projects. The FESTA methodology is not a rigid one; it is adaptable and kept alive by exchanging experiences and lessons learned from projects using the methodology. For automation studies, a common methodology is important, as stakeholders are not only interested in the findings of individual projects, but also in gaining knowledge about the wider impacts of automation.

The FESTA methodology puts a strong emphasis on defining research questions and hypotheses, inspired by the traditional impact areas of vehicle ICT: safety, mobility, environment and efficiency. With new functions enabling new mobility concepts and services, impact areas become wider (e.g. land use). Covering them requires considering multiple research questions of interest for stakeholders. New research questions create new information needs that may not be covered by current statistics.

The Knowledge Base for CCAM, maintained by the ARCADE consortium, continues to gather information on lessons learned from automation projects and on innovative methods and tools (CAD Knowledge Base).

In the coming years new steps need to be taken. As became clear from the gap analysis, there is a need for innovative and agile approaches as well as for practical solutions for the challenges CCAM evaluation poses. In order to be able to investigate the potential impacts of widespread CCAM deployment, results from multiple projects need to contribute to answering questions about the future of our transport system. We cannot predict the future, but we must strive to identify potential benefits and threats to society and to answer questions from industry and public authorities, and of citizens who will see their mobility options and the liveability of their environment change. For this a common evaluation method is crucial, ensuring that CCAM projects all contribute to a better understanding of the potential impacts, sharing common approaches and making results comparable.

In the coming years, a new European project will work on the development of a Common Evaluation Methodology (CEM) for CCAM, funded under the Horizon Europe Framework Programme. The FAME project (Framework for coordination of Automated Mobility in Europe) started in July 2022. The mission of FAME supports the commitment of the European Commission and the CCAM Partnership to provide a long-term coordination framework for research and innovation and large-scale testing and evaluation activities in Europe. The creation of a stakeholder-validated European framework for testing on public roads, including a CCAM Test Data Space (TDS) and the CEM, and the exchange of knowledge on CCAM activities will improve cooperation, consensus building and data sharing. Using FESTA and the new CEM in CCAM evaluation will support comparability and complementarity of the results of all the testing and evaluation activities.

Acknowledgements

The ARACDE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824251.

References

ARCADE EU support action. https://ec.europa.eu/inea/en/node/12589 & https://www.connectedautomateddriving.eu/about/arcade/

ARCADE, 2021. 4th Stakeholder workshop on Common Evaluation Methodology. Available at: https://www.connectedautomateddriving.eu/4th-stakeholder-workshop-on-common-evaluation-methodology/

Aittoniemi, E., Barnard, Y., Harrison, G., De Klein, D., Kolarova, V., Lehtonen, E., Malin, F., Naendrup-Poell, L., Rama, P. & Touliou, K. 2020. Autopilot (Automated Driving Progressed By Internet Of Things). D.4.6: Quality Of Life Impact Assessment. Available at: https://autopilot-project.eu/wp-content/uploads/sites/3/2020/09/AUTOPILOT-D4.6-Quality-of-Life-Impact-Assessment-v1.1.pdf

Barnard Y., Koskinen S., Wilmink I., Sanz L., Mesones A., 2021. Micro-FESTA for CCAM pilot projects. Available at: https://www.connectedautomateddriving.eu/micro-festa-v2/

Barnard, Y., Innamaa, S., Koskinen, S., Gellerman, H., Svanberg, E., Chen, H. 2016. Methodology for Field Operational Tests of Automated Vehicles. In: Rafalski, L., Zofka, A. (eds). Part of special issue for Transport Research Arena TRA2016. Transportation Research Procedia, Volume 14, Pages 2188-2196.

CAD Knowledge Base. https://www.connectedautomateddriving.eu/

CAD Knowledge Base Methodology. https://www.connectedautomateddriving.eu/methodology/

CARTRE EU support action. https://trimis.ec.europa.eu/project/coordination-automated-road-transport-deployment-europe/

CCAM partnership. https://www.ccam.eu/what-is-ccam/ccam-partnership/

FESTA Handbook Version 8, 2021. Available at https://www.connectedautomateddriving.eu/methodology/festa/

FOT-Net Data and CARTRE EU support actions. 2019. Data Sharing Framework Version 1.1. Available at: https://www.connectedautomateddriving.eu/wp-content/uploads/2021/09/Data-Sharing-Framework-v1.1-final.pdf

Giro, C., Barnard, Y., Page, Y., Amelink, M., Sanz, L., de Kort, A., Innamaa, S., Koskinen, S., Lenz, O. 2019. Society thematic areas: challenges and scenario. ARCADE Deliverable D3.7. Available at: https://knowledge-base.connectedautomateddriving.eu/wp-content/uploads/2020/04/EU-CSA-ARCADE_20191122_WP3_D3.7_Society_ChallengesScenarios_V2.0.pdf

Harrison, G., Shepherd, S. & Chen, H. 2021. Modelling Uptake Sensitivities of Connected and Automated Vehicle Technologies. International Journal of System Dynamics Applications (IJSDA), 10.

Innamaa S., Smith S., Barnard Y., Rainville L., Rakoff H., Horiguchi R., Gellerman H., 2018. Trilateral impact assessment framework for automation in road transportation, version 2.0. Trilateral impact assessment sub-group for ART. Available at: https://www.connectedautomateddriving.eu/methodology/impact-assessment-frameworks/

Innamaa, S., Louw, T., Merat, N., Torrao, G., & Aittoniemi, E. 2020. Applying the FESTA methodology to automated driving pilots. In Proceedings of TRA2020, the 8th Transport Research Arena, Helsinki, Finland. Available at: https://cris.vtt.fi/ws/portalfiles/portal/33906903/Applying_the_FESTA.pdf

Kalisvaart, S. et al. Proceedings of workshop on Common Evaluation Methodology. 2021. Available at: https://knowledge-base.connectedautomateddriving.eu/wp-content/uploads/2021/03/EU-CSA-ARCADE_Proceedings_Workshop_CEM_20201123.pdf

Rakoff, H., S. Smith, S. Innamaa, Y. Barnard, G. Harrison & Shaw. J. 2020. Building feedback into modelling impacts of automated vehicles: Developing a consensus model and quantitative tool. In Proceedings of TRA2020, the 8th Transport Research Arena. Helsinki, Finland. Available at: https://rosap.ntl.bts.gov/view/dot/48969