CROSS-DOMAIN SAFETY ASSURANCE

for Automated Transport Systems

CATAPULT High Value Manufacturing
UK Research and Innovation
WMG THE UNIVERSITY OF WARWICK
warwick.ac.uk/wmg

Table of Contents

DISCLAIMER

Contents of this report do not represent the position or policy of the Department for Transport or any of its agencies.

Executive Summary

The economic potential of the global automated transport ecosystem is projected to reach over £750 billion by 2035, with a UK market share of approximately 6% representing £42 billion and creating up to 38,000 new jobs[1].

However, safety remains the biggest challenge for commercialisation of automated transport systems (ATS)[2,3]. Safety and the corresponding perceived safety of ATS technology has a correlation with the development of trust and acceptance in the technology[4]. The universal nature of the impact of perceived safety on public acceptance of automated technology has been demonstrated by a variety of studies in the UK and internationally[5,6].

The current safety assurance frameworks in each of the transport domains (land, air and marine) have largely evolved within the constrain of their respective domain. This has resulted in a different set of strengths and weaknesses for each of them, despite having similar challenges. Furthermore, these differences limit the potential for cross-domain utilisation of tools, methods and skills. The same is true for the qualification processes for virtual test environments (VTE) which have become an integral part of the safety assurance process. This includes there being no common taxonomy or common understanding of the various stages involved in the assurance process, and little coordination between various programmes on proving safety and security of ATS and government initiatives across the transport domains (land, air and marine). With ongoing cooperation, there is an enormous opportunity to take best practices and learnings to create capability that provides cross domain skills with specific domain considerations. This report sets out a course of action and recommendations to address these issues.

In order to realise the huge commercial potential of ATS, it is essential that the ATS ecosystem (including the government) continues to evolve with technology in answering the safety challenges that technology presents.

If the ATS ecosystem stakeholders are serious about the safe introduction of ATS in society, then all stakeholders in the ATS ecosystem (technology developers, manufacturers, fleet service providers, regulators, policy makers, government etc.), should work together and focus on:

- Setting as high a standard of safety assurance as is appropriate to the domain
- Showing that deployed ATSs are safe, rather than competing and using safety claims as a unique selling point for individual brands.

The UK, with its history of a strong collaborative approach between industry, academia and government to new technology development, could demonstrate global leadership in both these areas by delivering these outcomes early and with high quality, and using them as exemplars of best practice.

For ATS technology, it is critical that methods and data related to safety are shared across the ecosystem. There are already case studies in some transport domains where industry (in the UK and internationally) collaborates in this way. For example, the aviation domain has a strong culture of data sharing when it comes to safety. Ongoing discussions at the United Nations Economic Commission for Europe (UNECE) forums reflect a similar approach for the automotive domain too.

When it comes to safety assurance of automated transport systems, we suggest the need to undertake two key steps:

- Establish the safety level of the ATS
- Communicate the safety level of ATS

To establish the safety level of the ATS, it is best to establish societal concerns and tolerance of undesirable outcomes. Once these are agreed, it is possible to define the safety targets/requirements across the system. These considerations will enable the development of a correct and complete set of requirements that will drive the assurance process. A key element of these requirements will be the elicitation of the operating conditions. This means the description of the Operational Design Domain (ODD) of the ATS in an objective manner. The ODD defines the conditions the ATS can handle safely when deployed in the world, and importantly those that that it cannot and must be appropriately prevented. An ODD includes geographic locations, weather conditions, actors in the locations etc. Using the ODD approach allows for a manageable, consistent, and scalable assurance process within each domain. This makes the ODD approach (i.e., the process) interoperable between domains.

An ATS will encounter many different situations in deployment. Until recently, the principal method for the ATS to experience these situations was to perform "real-world testing". Over the past few years, however it has been recognised that this is impractical and that there is a need to identify, capture and store these situations in a more efficient and systematic way. These captured situations are stored as scenarios and can then be used as part of a coordinated testing programme including the use of "virtual test environments (VTE)" (i.e., simulation). If the evidence generated from VTE is to be used as part of the performance and safety argument of the ATS' capability, then it is imperative that the virtual test environments are trusted through their own verification and validation.

The development of a safety assurance process will include a diverse set of stakeholders. They will need to work together and understand each other. This will be facilitated by a standardised set of taxonomies and definitions for each aspect of the safety assurance process.

Research[4] has shown that people's trust and acceptance of automation is influenced by the perceived safety of the technology. The fundamental reason for this is to alleviate concerns to the point that users are suitably confident in the service provider. Therefore, accurately communicating the safety levels of the technology to users in an accessible manner is the first step to establish confidence and will lead to user acceptance.

Establishing and communicating the various aspects associated with the safety of ATS are common challenges for each of the transport domains (land, air and marine).

[1] Connected and Automated Vehicles: market forecast 2020, Connected Places Catapult (2021). https://www.gov.uk/government/publications/connected-and-automated-vehicles-market-forecast-2020

[2] Koopman, P. and Wagner, M., 2017. Autonomous vehicle safety: An interdisciplinary challenge. IEEE Intelligent Transportation Systems Magazine, 9(1), pp.90-96.

[3] Rezaei, A. and Caulfield, B., 2020. Examining public acceptance of autonomous mobility. Travel behaviour and society, 21, pp.235-246.

[4] Khastgir, S., Birrell, S., Dhadyalla, G., & Jennings, P. (2018). Calibrating trust through knowledge: Introducing the concept of informed safety for automation in vehicles. Transportation research part C: emerging technologies, 96, 290-303.

[5] Stilgoe, J., 2021. How can we know a self-driving car is safe?. Ethics and Information Technology, 23(4), pp.635-647.

[6] Körber, M., Baseler, E. and Bengler, K., 2018. Introduction matters: Manipulating trust in automation and reliance in automated driving. Applied ergonomics, 66, pp.18-31.

Our Key Recommendations Include:

Establishing Safety Levels

- **Develop a common safety assurance framework for automated transport systems in land, air and marine.**

  This should take account of existing systems, building on them, and be based on the concept of Operational Design Domain (ODD) (i.e., the limit of operating conditions) which allows for a manageable, consistent, and scalable assurance process within each domain, and interoperability between domains.

- **Create a qualification process for Virtual Test Environments (VTE) to enable trust in evidence from virtual testing.**

  The use of Virtual Test Environments (VTE) will be essential for the safety assurance of ATS (in each domain). The qualification process needs to ensure that all the virtualised components of the test environment are appropriately validated.

- **Create standardised taxonomies and definitions for each element in the assurance process.**

  The wide variety of stakeholders involved in the safety assurance process demands a common understanding of the various elements involved in the assurance process (i.e., ODD, test scenarios, safety metrics and thresholds). This requires clear commonalities where possible (e.g., levels of automation etc.).

- **Coordinate research and government programmes and government initiatives across all domains.**

  There is a need to review, prove and where possible consolidate the approach for the safety of automated transport systems across the various transport domains. This is facilitated by the exchange of knowledge and skills between the domains and will make commercial and economic exploitation more efficient.

Communicating Safety

- **Create a common set of principles to communicate how safety is assured to all stakeholders in each transport domain.**

  The message will need to be tailored to the relevant audience and the content of the communication will need to vary depending on the type of stakeholder (e.g., public, developers, regulators, insurers).

- **Encourage independent organisations to take a proactive role in communicating safety of ATS.**

  The credibility and independence of the organisations and people communicating safety is key. Establishing trust and credibility in the messaging by reputable organisations and individuals will help to grow trust in the technology.

1. Background

Transport and mobility are a cornerstone of our society and our economy, accounting for over £109 billion of UK economic added value in 2019[7].

When transport and mobility improve, society improves, and the economy grows. It enables the transportation of medicines, food, and products, and the movement of people for work, family, and play. Transport is also in the spotlight as part of the decarbonisation agenda and improvements in transport utilisation and efficiency could help to meet targets associated with the climate change agenda.

Automation has changed our world. The dictionary definition is the use of machines or computers that can operate without human control, but that disguises its impacts. Most people would think of automation in the realm of factories – where jobs that are simple, repetitive and probably boring activities have been replaced with machines doing those jobs. That enabled them to be carried out at much greater scale and satisfy the growing markets for products. A by-product of this change was that the consistency and quality of products improved because machines do not lose focus, get bored or otherwise get distracted from their allotted task. Automation can lead to some "human jobs" becoming redundant, but normally a new set of higher skilled higher value jobs are created to develop, manage, and maintain these systems. Thus, automation can also enable the creation of better jobs for society.

There are additional examples where automation has positively impacted society. The humble washing machine is one of many domestic automated machines that have freed up society from repetitive and time-consuming tasks, and have improved our quality of life.

The reason for most travel is to allow passengers or goods to reach their destination, rather than to undertake the journey itself. Many aspects of travel are made easier by the use of some level of automation, e.g., automated ticketing, computerised scheduling, driver assistance, and automated flying.

Automation in cases such as in the factory or domestic device, can be achieved by defining the rules that control the process. If, the set of rules and the possible number of decisions is limited and can be identified, then this can be implemented as a set of control instructions. This approach can be applied, and it works well, if you are on a train with limited track changes, or on a plane in clear skies, or a ship crossing an open ocean. Unfortunately, in many situations, outside those ideal situations, the control system needs something more than simple rule-based automation – it needs autonomy. Autonomy has several definitions, but all include a degree of self-determination. Rather than simply following the rules, an automated system can make decisions even when the circumstances fall outside the strict rule of definitions.

A good example of when this happens is transport, i.e., the ubiquitous car, aeroplanes or ships. A motorway journey, which is probably closest to being controllable by rules alone, quickly strays outside simple control territory with the presence of other cars, the inevitable weather, and many other factors. The same is true for flights which might face turbulent weather and ships which have to sail in varying sea states. In an urban driving case, the roads are rarely straight, the traffic flow is constantly interrupted by junctions with other roads and the space is shared with cyclists, pedestrians and other objects. Automation is not easy!

Since it is not easy, the question of whether it is worth it arises.

Do people want to be able to get into a small personal transport and spend the journey working or even watching the scenery rather than focusing on driving or operating the system?

Part of the answer is in one of the other attractions of automation, i.e., consistency. How would that manifest itself for an automated vehicle (land, air or marine)? Perhaps the answer is safety.

The majority of transport accidents have human error as a contributing factor. If it can be shown that automated transport systems are safer and less likely to be involved in an accident, then that should help build a compelling argument for ATS.

[7] UKRI: Innovate UK, 2021, UK TRANSPORT VISION 2050: investing in the future of mobility.

To achieve the vision of a safe automated transport system, it will be necessary to merge the physical worlds of mobility and infrastructure with the digital world of data, software, simulation, and Artificial Intelligence. The resulting system will be complex and produce several intertwined challenges in many areas including and possibly especially safety.

So, proving the safety of an ATS is a challenge, but one that needs to be met if its benefits are to be delivered. The workshops that have helped to produce this report have shown that there has been a siloed approach to the safety assurance of automated transport solutions across the different transport domains of land, air and marine. However, those workshops also identified that there was a high degree of commonality in the challenges that each domain is facing.

This report recommends that the challenges and issues faced by the transport sector need to be addressed across all transport domains (land, air and marine) to enable maximum potential and opportunity to be realised.

To maximise the benefits, will require integration and collaboration between domains, manufacturers, infrastructure, transport service providers and regulators. To achieve a credible, robust and effective safety methodology for ATS (land, air and marine), it must set high standards for safety assurance and also enable safety to be a collaborative, not competitive, endeavour. If the methodology can deliver on these key principles and maximise cross-domain commonalities, then it will accelerate safety, trust and acceptance to maintain pace with technology's potential.

The pivotal role of safety assurance in ATS is widely recognised in Governmental policy (e.g., 'Connected & Automated Mobility 2025[8], The National AI strategy[9]). Furthermore, the recent Innovate UK Transport Vision 2050[10] recognised the need to exploit opportunities 'across mobility modes'. This report presents an approach to identify similar challenges each domain is facing with the introduction of autonomy such as defining safe enough, proving safe enough, regulatory frameworks and a common safety assurance methodology. This approach can be applied to each domain separately but the underlying commonality will allow all the domains to learn and accelerate more rapidly because they can benefit from the knowledge gained across all the domains.

Addressing this challenge will not only enhance the UK's transport and mobility offerings, but it will also build industrial capability and value. The UK's own automated transport strategy projects a global £750 billion connected and automated transport systems market by 2035, with a UK market share of approximately 6% representing £42 billion and creating up to 38,000 new jobs[11]. Addressing the safety challenge will have ramifications for UK technology developers, insurance providers, cross domain integrators, mobility operators, data collection and analysis, tier 1 and 2 supply chain (i.e., sensors, software stack, simulation and testing), and many more associated providers. The benefits of a cross-domain approach to policy, standards and regulation will focus technological development and act as an enabler to UK industry and offer global opportunities.

The UK has one of the world's leading research outputs, underpinned by a network of world-class universities. It is imperative we convert this intellectual capital into world leading policy, products, services, and deliver positive societal and economic outcomes. This report shows how this can be done with a cross-domain approach to the safety assurance of ATS in land, air and marine domains. The report captures the collective intelligence of academia, industry regulators and Government applied to the problem of an open, cross-domain approach to safety assurance.

[8] HM Government, 2022. Connected & Automated Mobility 2025: Realising the benefits of self-driving vehicles in the UK.
[9] HM Government, 2021, National AI Strategy.
[10] UKRI: Innovate UK, 2021, UK TRANSPORT VISION 2050: investing in the future of mobility.
[11] HM Government, 2022. Connected & Automated Mobility 2025: Realising the benefits of self-driving vehicles in the UK.

2. Motivation:
Why did we do this work?

Automated Transport Systems (ATS) in each of the transport domains (land, air and marine) will require to be proven safe before and during their commercial deployment.

However, each transport domain will have a variety of use cases for commercialisation. Some may involve carrying passengers from one location to another on a pre-defined route, others may include transporting logistics in a very constrained environment. There may be many other uses cases too.

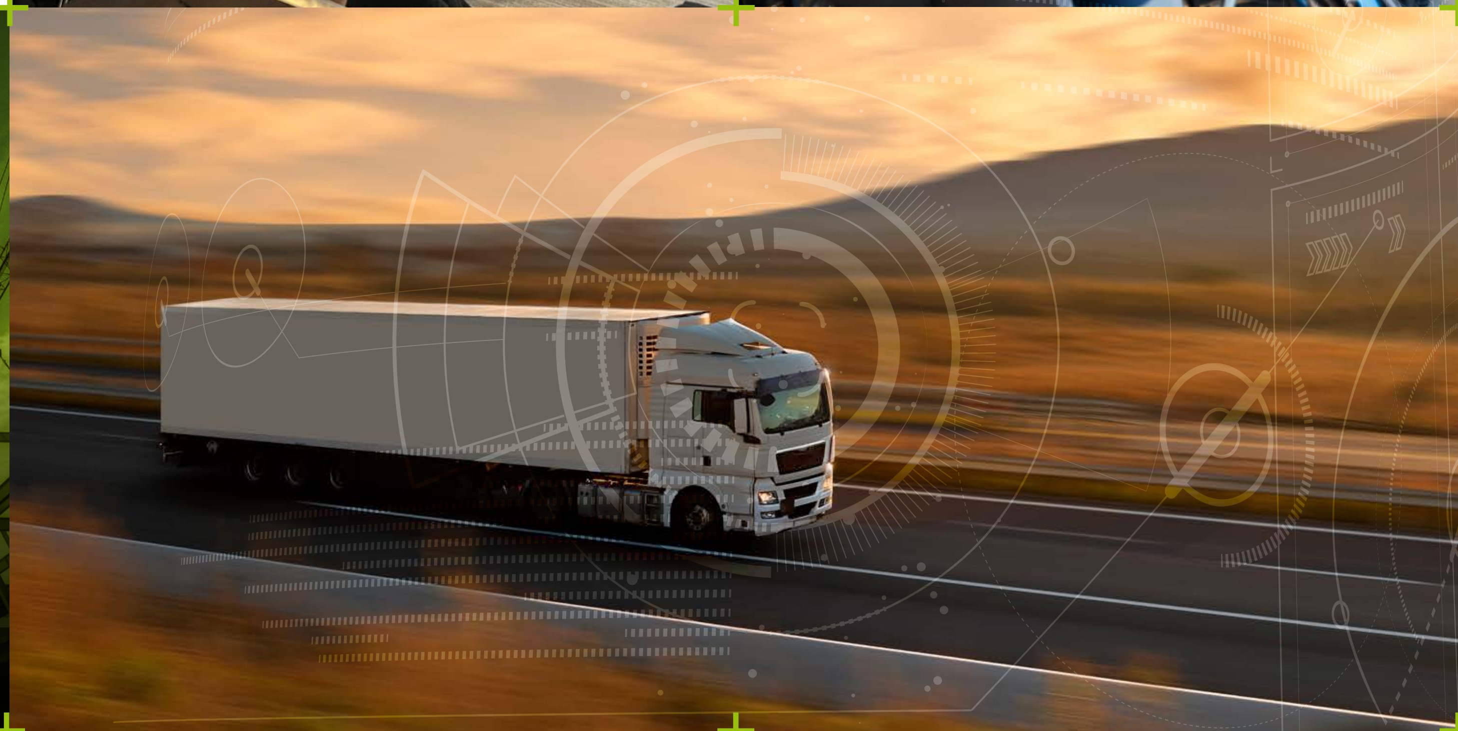
Interestingly, the research questions to prove their safety for all the use cases are fundamentally similar across the transport domains. **Some of these include:**

Defining "safety"

Proving "safety"

Communicating "safety"

In order to have efficient safety assurance processes, it is important to create a methodology for answering the above which is both scalable and pragmatic while maintaining high benchmarks for safety. This will enable the re-use of tools, processes and investment across the various transport domains in a holistic manner.



3. The Methodology:
How did we get here?

In March 2022, WMG organised an event to evaluate the opportunity for a cross-domain safety assurance framework.

This was borne of previous disparate discussions that suggested an opportunity to enhance the safety in all of the domains by sharing experience and systems.

The day went well, with many productive discussions among the 100 people present from various modes of transport. One of the themes from the day was that when it comes to safety assurance of automated transport systems, not only do we need to establish the safety of the system, but also communicate the safety to the various stakeholders of the ecosystem, including society.

As a result of the discussions on the day, three areas were identified to take the work forward:

- Developing a safety assurance framework that would combine the learning and approaches of the separate domains to allow cross sharing.
- Exploring how to incorporate virtual test environments into the methodology for validating safety of the systems.
- Understanding the difference between measured safety performance and the perceived safety of a system.

For each of these areas, WMG undertook to set up and manage working parties to develop and agree on ideas to answer these challenges. As part of the working party discussions, working party 1 focussed on answering the question of *safety assurance framework*, working party 2 focussed on *virtual testing* and working party 3 focussed on *perceived safety*.

For the latter part of 2022 and into early 2023, each of these working parties met three times, initially with mainly regulatory body participants but gradually widening participation to include developers and potential users of automated transport systems in the three domains (land, air and marine).

About half-way through the process, the discussions uncovered that "remote operation" was analogous to automated transport, and learnings could be transferred between these systems and fully automated systems.

An additional workshop was therefore convened to explore this area, and the similarities and differences explored and tested. There were over 200 person days of discussions over 10 working party meetings and participants from more than 30 national and international organisations at the working party meetings, and it is these conversations and ideas from these workshops which led to the content and recommendations of this report.

Figure 1 Timelines of various cross-domain working party meetings

March 2022
Concept launch

Working Party
Formation & Stakeholder mapping

Oct 2022
Working Party 1-2-3
→
Nov 2022
Working Party 1-2
→
Dec 2022
Working Party 3
↓
Jan 2023
Remote Operation workshop
←
Jan 2023
Working Party 1-2 (final)
←
Feb 2023
Working Party 3 (final)

March 2023
Launch of Cross Domain report with findings

4. The Findings
Understanding Assurance

Out of the three themes identified in the March 2022 workshop, two of them focussed on understanding safety assurance and the tools required for this. Here we discuss the findings of these two working parties.

From a first principles perspective, irrespective of the transport domain (land, air and marine) or their corresponding use cases, assurance for automated transport systems needs to be undertaken at three separate levels:

- Performance assurance
- Process assurance
- Usage assurance

Performance assurance means assuring the performance of the automated transport system (ATS). Process assurance means assuring the process of development and maintenance of the ATS. Usage assurance means assuring the capability of the organisation and the workforce to implement and use the developed processes.

4.1. Performance Assurance

It is important to establish that performance assurance needs to be undertaken both before deployment and during deployment of the ATS. This is essential as the ATS may receive a software upgrade (over the air) or have degraded performance due to gradual or sudden failures.

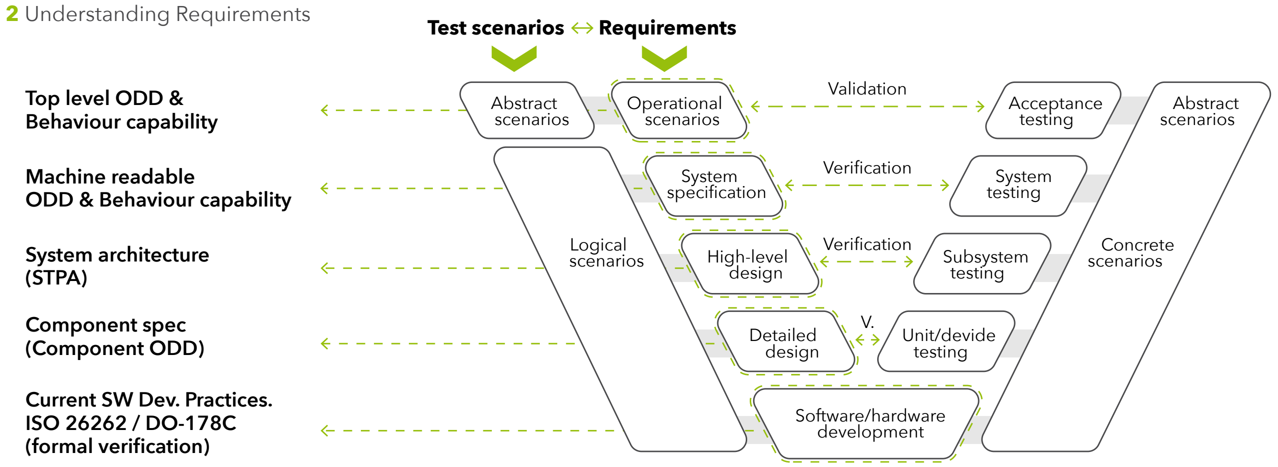
4.1.1. Three pillars of Performance Assurance

Once the requirements have been established, the performance assurance comprises of three building blocks. These include: 1) Test scenarios 2) Test environment, 3) Safety argument.

A test scenario illustrates the situations an automated transport system will experience during its deployment in the real world. An understanding of these situations is key to designing a safe automated transport system. A test environment is the platform in which the ATS undergoes testing. A safety argument provides the link between safety evidence and the safety claim (i.e., the system is safe to use). Having executed the test scenarios in the test environments, we need to analyse the results and decide if they indicate that the behaviour of the system is safe.

The first step towards performance assurance is to develop an accurate and clear understanding of requirements. One of the key parts of developing requirements for an ATS is to understand its operating conditions and behaviour capabilities. The operating condition of an ATS which is a design specification is also known as the Operational Design Domain (ODD)[12]. An ODD definition will include all static, dynamic and environmental attributes like weather, connectivity etc. To ensure completeness of requirements from an ODD perspective, it is essential to capture the wide variety of actors and their diverse characteristics in an ODD definition. This should include characteristics like type of actor, skin colour, disabilities etc. An ODD definition doesn't define the behaviour capabilities or the desired behaviour of the system. As a result, the ODD and the behaviour capabilities definition together comprise the system concept for an ATS **[fig 2]**.

Figure 2 Understanding Requirements



[12] SAE J3106: https://www.sae.org/standards/content/j3016_202104/

4.1.2. Understanding ODD and TOD

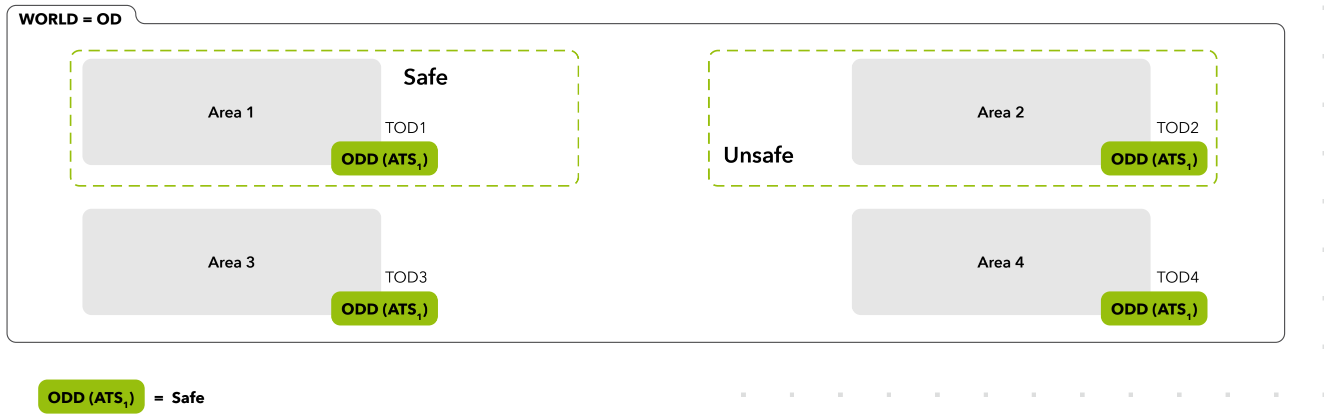
ISO/FDIS 34503 is an upcoming international standard that provides a taxonomy for ODD definition for the land domain. It also introduces the concept of Operational Domain (OD) and Target Operational Domain (TOD)[13]. The concept of TOD together with ODD forms the bedrock of the performance assurance process.

TOD refers to the set of operating conditions in which an ATS will be expected to operate.

This may include environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain road users etc.

The key difference between ODD and TOD is that ODD expresses a specification of the ATS, whereas TOD is a description/specification of an environment in which an ATS will be expected to operate **[fig 3]**. An ATS may be proven to be safe in its ODD [fig3], but that doesn't mean that will also be safe when deployed in different TODs (e.g., Area 1 and Area 2). Depending on the overlap between the ODD of the ATS_1, and TOD_1 and TOD2, the ATS_1 may be safe in Area 1 while being unsafe in Area 2.

Figure 3 ODD and Target Operation Domain (TOD)



4.1.3. Why is ODD important?

As part of scenario-based testing of an ATS, it is important to consider what kind of scenarios an ATS would experience. The set of scenarios an ATS will encounter are directly related to the area of deployment of the ATS. For example, the scenarios a self-driving vehicle will experience in central London are different from those it would experience in Leeds to those in Phoenix, US. Similarly, an automated vessel will experience different scenarios while docking in the port of Tokyo to those in Singapore to those in Plymouth which will also be very different from those in open seas. Aerial drones will have a similar experience. Aerial drones flying up to 1000 ft altitude to those flying in controlled airspace to those flying over residential areas will face different scenarios. Therefore, scenarios used as part of performance assurance will need to be directly related to the ODD of the ATS.

Additionally, as ODD influences the scenarios to be tested, it therefore impacts the fidelity requirements for the Virtual Test Environment (VTE) used for virtual testing which will need to be able to execute those scenarios. For example, if rainfall is included within the ODD of an ATS, the VTE needs to be able to represent the effect of rainfall on sensors as part of the VTE modelling. It is important for the scenarios to show appropriate coverage of the ODD meaning that the coverage metrics are also directly influenced by the ODD.

Thus, each aspect of the performance assurance has a direct correlation with the ODD of the ATS. Using the Operational Design Domain approach allows for a manageable, consistent and scalable assurance process within each domain, and allows interoperability between domains.

[13] ISO/FDIS 34503: Road Vehicles – Test scenarios for automated driving systems – Specification for operational design domain. https://www.iso.org/standard/78952.html

4.1.4. ODD based scalable "performance assurance" framework

Each of these three pillars of performance assurance can be further divided into more concrete steps [fig 4]. The Test Scenario pillar has three phases: 1) Create 2) Format and 3) Store. The Test Environment pillar has two phases: 1) Plan and 2) Execute. The Safety argument pillar which has two phases 1) Analyse and 2) Decide.

In order to create a scalable performance assurance framework, the flow:

create ❯ format ❯ store ❯ plan ❯ execute ❯ analyse ❯ decide

needs to be underpinned by the ODD and the behaviour capability definition. Irrespective of the transport domain (land, air and marine) or their use case or the complexity of their ODD, each ATS will always need to follow this flow for performance assurance.

The scenarios pillar identifies the content of the scenario (create), formats them in a standardised language (format) and stores them in a scenario library enabling reuse of the scenario (store).

Once the test scenarios have been identified, we need such test environments to execute the test scenarios. An ATS will experience a large number of scenarios during the course of its deployment lifecycle. Developers would be expected to test against such scenarios as part of the development and assurance process of the ATS. However, it wouldn't be plausible to execute such a large number of scenarios in the real-world by driving/ flying/ sailing. The use of Virtual Test Environments (i.e., simulation) to execute these tests becomes imperative in this case.

A continuum of test environments from the entirely simulated or virtual, through physical test beds, to real-world settings, will provide the range of options at a reasonable speed to ensure the safety of automated transport for land, air and marine. While the use of VTEs as part of the testing continuum is obvious, in order to trust the evidence from test executions in a VTE requires the VTE to be qualified, i.e., proven to be representative of the real world.

In order to ensure that an ATS is performance assured, not only do we need to test the system against its ODD definition, but also in conditions outside its ODD and in its TOD. This introduces the concept of a **system safety** case and an **operational safety** case. A system safety case provides the argument that the ATS is safe in its ODD. **On the other hand, an operational safety case provides the argument that the ATS is safe when deployed in its TOD (i.e., specific part of the world)**. An operational safety case will change for each deployment.

A system safety case will be constant for all deployments unless system modifications are made, at which instant a new assessment will be required.

In order to structure the safety case there are a number of formats that could be used[14]. The focus of this report is on the creation of the content that goes into a safety case (of any format).

Figure 4 ODD based Scalable Safety Case Framework

[14] Kelly, T., & Weaver, R. (2004, July). The goal structuring notation–a safety argument notation. In Proceedings of the dependable systems and networks 2004 workshop on assurance cases (Vol. 6).

4.1.5. Test Scenarios

For automated systems, research has shown that it is more important to test *"how a system fails"* as compared to *"how a system works"*[15]. As a result, we propose a hazard-based testing concept where the focus is to use scenarios to expose failures in the ATS. We propose a hybrid methodology of knowledge-based scenario generation and data-based scenario generation **[fig 5]**. The test scenarios pillar has three distinct phases: 1) Create, 2) Format and, 3) Store.

Create

Data-based scenario generation leverages real-world data from accidents, insurance records, or data from real-world trials to understand trends in the accidents and near-miss scenarios. Knowledge-based scenario generation focusses on analysing the system design of the ATS to understand how the design of the system could lead to a failure. Both data-based and knowledge-based methods are focussed on revealing failures. It is also possible to use multiple synthetic data generated scenarios (generated in virtual test environments using a single ground truth data, i.e., real world data) as a means of scenario searching and expansion.

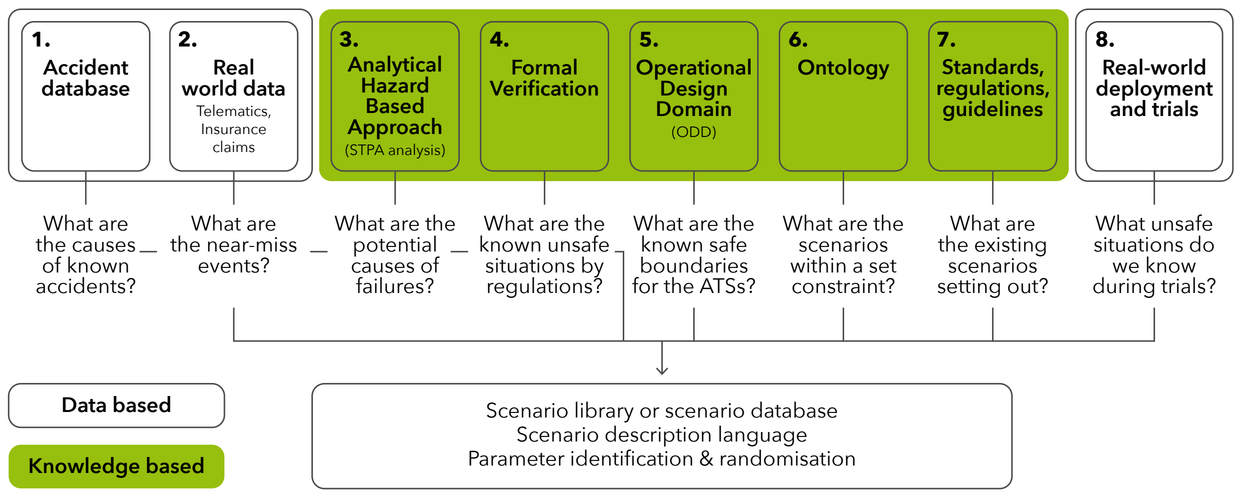
Format

The content of the scenario generated in the "create" pillar will need to be communicated and shared with a diverse set of users (e.g., ATS developers, suppliers, regulators, insurers, public etc.). Thus, there is a need for a common format illustrating the scenario content to ensure a common understanding of the test scenarios. The "format" pillar provides that an adequate scenario description format is used to represent the content and enable its sharing and execution. It is essential that this developed scenario description format meets the requirements of various end-users who sit on different positions along the product development cycle.

Store

Once the scenarios have been identified and formatted in a standardised language, they will need to be stored in a scenario library. This will be necessary due to the large number of scenarios required for performance assurance as well as to prevent the need to reinvent the wheel for every performance assurance activity.

Figure 5 Various scenario generation methods



[15] Khastgir, S., Birrell, S., Dhadyalla, G., & Jennings, P. (2018). The science of testing: An automotive perspective (No. 2018-01-1070). SAE Technical Paper.

4.1.6. Test Environment

Having identified, formatted, and stored the scenarios in a scenario library, the next step involves the execution of the scenarios in various test environments. In this report we focus primarily on the Virtual Test Environment (VTE) as it was one of the focus themes identified in the March 2022 workshop. VTEs have been used to varied levels in the land, air and marine transport domains. While the aviation sector has made extensive use of VTE (i.e., simulation), its use for land and marine has been somewhat limited when environment modelling is considered. It is worth noting that while the aviation sector has had widespread use of simulation and virtual test environments, they have primarily been used for pilot training and not to assure the performance of automated systems.

Taking a systems approach to qualification of virtual test environments (VTE), it needs to be undertaken at three separate levels (similar to assurance of the ATS): 1) Performance, 2) Process, 3)[16] Usage. Figure 7 illustrates the qualification framework for VTE with these three aspects. It is important to highlight that a VTE shall be considered as a qualified VTE, if and only if all three levels of qualification have been met[16]. Performance qualification considers the performance of the VTE which is used for ATS testing, i.e., the combination of the environment model and the sensor model. This is done as the output of the sensor model (having perceived the environment), is the input to the ATS control system **[fig 6]**. Process qualification (via audit) focusses on the process of development of the VTE. Usage qualification focusses on qualification (via audit) of the capability of the workforce to build (on the VTE developer side) the VTE and use (on the ATS developer side) the virtual test environment.

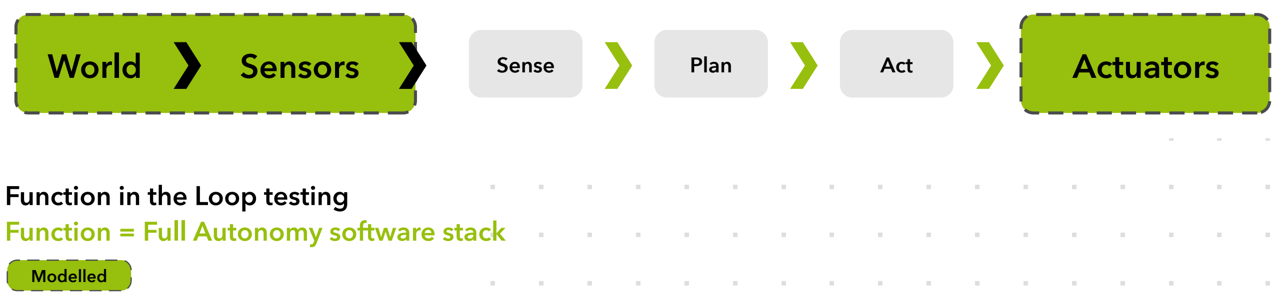
While various other tools and qualification approaches have been suggested in existing standards like ISO 26262[17], ISO 21934[18], MIL-STD-3022[19], NASA-STD-7009[20] and NAFEMS Engineering Simulation Quality Management Standard (ESQMS)[21], they either completely miss out on performance qualification of VTE or do not provide detailed guidance on performance qualification of the individual sensor/environment model combination.

Performance qualification of VTE

For performance qualification of the VTE, the focus should be on the comparison of raw sensor outputs from the real-world sensor in the real-world, and the virtual sensor in the virtual world, and the response of the actors (in real world and in VTE). This needs to be done by collecting data from deploying the ATS on the same path in the real and the virtual worlds. The comparison between the sensor outputs needs to be made for both static elements and dynamic elements in the ODD of the ATS. The concept of "performance qualification" is generic across sensor configurations (i.e., LiDAR, RADAR or camera, ultrasonics, SONAR sensors) and suggests that the comparison is done on the output of the real-sensor and the virtual sensor.

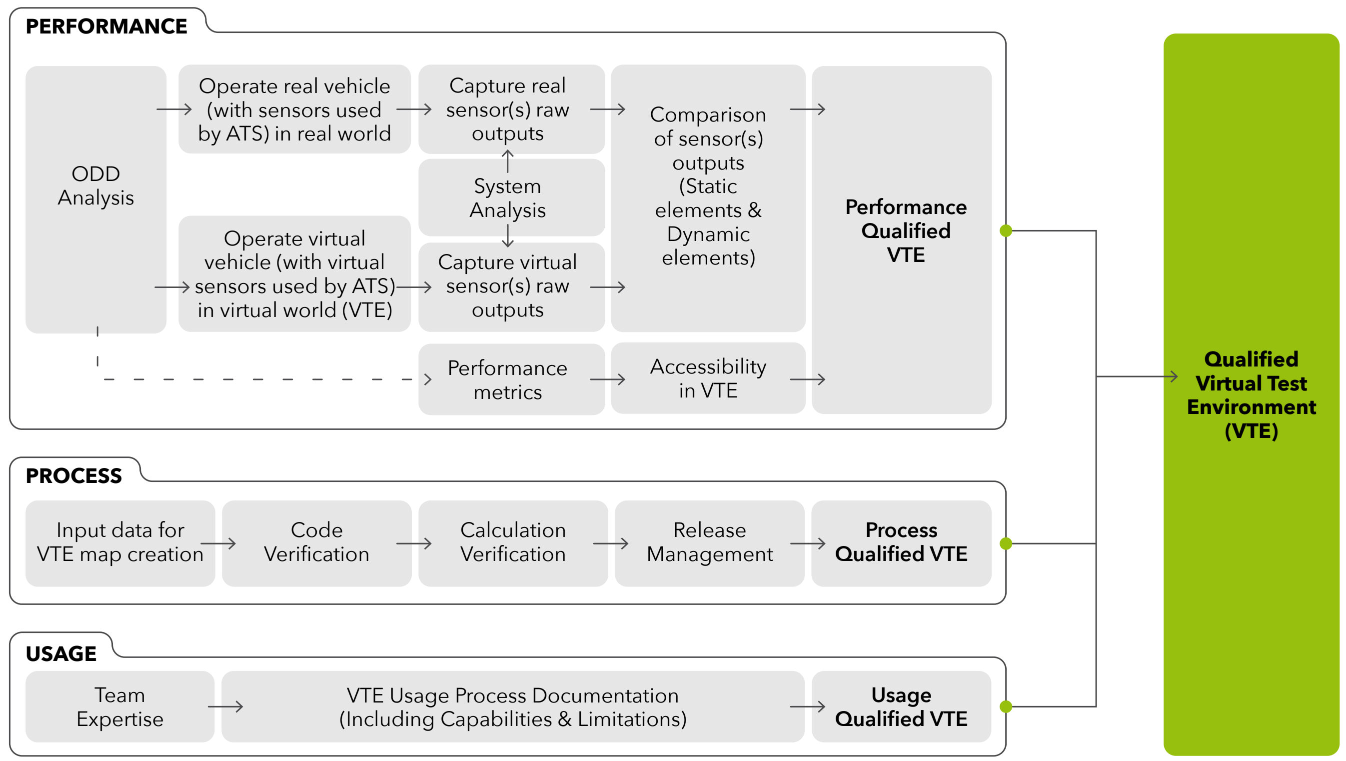
Similar to the performance qualification of the ATS, the first step in the qualification of performance **[fig 7]** of VTE involves the analysis of the ODD. This is essential to understand not only the type of static and dynamic elements needed for modelling in the VTE, but also the environmental conditions whose effect will need to be reflected in the VTE. This would include both modelling the environment in the VTE and the effect of the environment on virtual sensor model performance in the VTE. "Systems Analysis" involves the identification of the types of sensors that will need to be modelled virtually, a key aspect for the VTE performance.

Figure 6 Possible components to model in a VTE



[16] Introducing a Qualification Scheme for Virtual Test Environments for Driving Automation Systems. Traffic Injury Prevention. (in review).
[17] ISO 26262. Road Vehicles - Functional Safety. (2018).
[18] ISO 21934 (2021). Road Vehicles - Prospective safety performance assessment of pre-crash technology by virtual simulation.
[19] Department of Defense (US). (2008). MIL-STD-882 - Documentation of Verification, Validation and Accreditation (VV&A) for models and simulations. In Mil-Std-882E (Issue May).
[20] NASA Standard 7009a - Standard for models and simulations (2016).
[21] NAFEMS. (2020). Engineering Simulation Quality Management Standard. https://www.nafems.org/publications/resource_center/esqms-01/

Figure 7 Qualification scheme for Virtual Test Environments (VTE)



Every comparison run between the sensor outputs will involve the execution of a scenario (in the real world and the virtual world). As the scenario parameter range (i.e., logical scenario) can have a variety of values, it is also essential to perform a sensitivity analysis on the parameter values to establish if any parameter (and its value) has a higher than usual effect on the comparison of the sensor outputs[22].

As part of the performance qualification, it is imperative that the qualification is done in a manner that takes into account both the fidelity of the virtual environment (i.e., world) and the virtual sensor model by making the comparison at the level of the "sensed" virtual environment **[fig 6]**. This is key from a virtual testing perspective as the ATS responds to the sensed environment. Furthermore, the sensed sensor outputs need to be done in a coherent manner and a comparison of the coherency of real-sensor outputs and the virtual-sensor outputs needs to be evaluated.

Another aspect of qualification of performance of VTE involves the ability to provide access to the various performance metrics used to make an evaluation of the ATS performance. The performance metrics for an ATS will also be affected by the ODD of the ATS. It should be noted here that the performance of the VTE is different from performance of the ATS. Together with the comparison of the (real and virtual) sensor outputs, coherency of these outputs and the ability to access the performance metrics (or the variables associated with it to calculate the metric), the performance of a VTE can be qualified.

It is important to highlight that every VTE will have its "sweet spot" for operation, and it should be kept within these boundaries when used as part of assurance in order to have confidence in the results.

The working party findings suggest that this sweet spot would be dependent on ODD parameters, use of VTE, safety thresholds and confidence limits.

Process Qualification of VTE

Qualification of the development process of the VTE will include various stages and will generally need to be done via an audit process. This is similar to the tool qualification mentioned as per ISO 26262 – Part 8[23]. As the qualification of performance of the VTE has a significant dependence on the creation of the base map of the VTE in which the virtual testing is performed, one of the first steps in the development process qualification should involve the audit of the input data used for VTE map creation.

This generally is a combination of high-density sensor (e.g., LiDAR, camera, SONAR, RADAR) scan images to create a (photo-realistic) 3D environment.

The next stages in the process qualification involve the code verification, which is followed by calculation verification.

Code verification refers to checking if there exists any numerical or logical flaws in the virtual models in the VTE. Calculation verification refers to the estimation of the number errors[24]. Like any software package, a VTE will also undergo periodic maintenance and upgrades with new improved versions of the software released. Thus, the last stage of the development process qualification focusses on release management which requires traceability across software versions and various VTE development work products. This illustrates the need to shift from singular evidence to lots of evidence while auditing in order to judge the qualification of the VTE.

Usage Qualification of VTE

To have a qualified VTE for ATS testing, it is essential to take a systems approach to the qualification process. For organisations involved in the development or testing or approval of ATS using VTE, their biggest interest in the qualification process is in having an assurance that the correct results were obtained as part of the virtual testing of the ATS. Ultimately, the VTE platform will be used by humans for development testing, approval testing etc. Thus, the correctness of the results using a VTE is dependent on two aspects: 1) correctness of the tool (i.e., VTE) and 2) correct use of the tool by the user (i.e., human who is also part of the system). The qualification of VTE usage will generally be done by an audit process. This should not only include audit of the team's expertise but also of the documentation of the VTE (both from the VTE development perspective as well as ATS development perspective). ISO 26262-Part 2 also suggests that "persons involved in the execution of the safety lifecycle have a sufficient level of skills, competence and qualification corresponding to their responsibilities". As personnel performing testing of an ATS using the VTE are part of the safety lifecycle, the corresponding ISO 26262 requirement also applies here. Similar to a user using an ATS, any such documentation should clearly articulate both the capabilities and the limitations of the VTE, to enable informed decision making by the users[25]. Our proposed qualification scheme or VTE includes both the qualification of the tool and the user of the tool.

For a VTE to be deemed as qualified, it will need qualification of performance, development process and usage. In the absence of any of the three elements, a VTE cannot be considered to be a qualified VTE.

4.1.7. Safety evidence and safety argument

The safety evidence and safety argument pillar has two distinct phases: 1) Analyse and, 2) Decide. In the analyse phase, the pass / fail criteria for each of the test executions is defined. This involves the definition of the "safe behaviour" for the ATS which can be used as one of the pass / fail criteria.

To this end, considerable discussion around safety for ATS focusses on being as good as, or better than, human driven systems on land, air and marine. Behaviour of human beings in transport systems are governed by a set of rules (e.g. rules of road (The Highway Code), rules of air UK SERA and rules of sea (COLREGs)).

However, as human beings we are exceptionally good at handling unfamiliar situations by using our intuitive pattern matching ability, which enables us to safely handle the "edge case" situations - those which bear a resemblance to situations we have experienced before, but which are not exactly the same. A computer simulation only has these links if it has been programmed to, or if it has "learnt" directly from artificial intelligence training.

There is a need for codification of the "safe behaviour" definition for land, air and marine for ATSs.

This will involve deriving inspiration from the Highway Code (land), COLREGs (marine) and UK SERA (Standardised European Rules of the Air) (air). The land domain has taken an ODD based approach in codification of the rules of the road[26], and a similar approach may be taken for codification of rules of the air and rules of the sea.

In the decide phase, the coverage criteria for the suite of test scenarios are defined. We can then use the collected evidence to create the safety argument proving that the ATS is a safe system. This will entail comparing against "defined safe behaviour" and "defined coverage criteria".

22 Dona, R., & Ciuffo, B. (2022). Virtual Testing of Automated Driving Systems. A Survey on Validation Methods. IEEE Access, 10, 24349-24367.
23 ISO 26262. Road Vehicles - Functional Safety. (2018).
24 UNECE VMAD - SG 2 (Virtual testing). (2022). https://wiki.unece.org/pages/viewpage.action?pageId=117508578

25 Khastgir, S., Birrell, S., Dhadyalla, G., & Jennings, P. (2018). Calibrating trust through knowledge: Introducing the concept of informed safety for automation in vehicles. Transportation research part C: emerging technologies, 96, 290-303.
26 UK Rules of the Road proposal: https submission to UNECE FRAV ://wiki.unece.org/display/trans/FRAV+33rd+Session

4.2. Process Assurance

As performance assurance requires the execution of the test scenario in a test environment (virtual test environment, test track or real-world testing), it would be implausible to make an argument that we can assure the performance of the ATS in all possible scenarios from an approval process. This is due to the amount of time it would potentially take to execute billions of test scenarios. Even though the use of virtual test environments provide more flexibility in increasing the number of executions, we will never be able to execute all potential scenarios during approval.

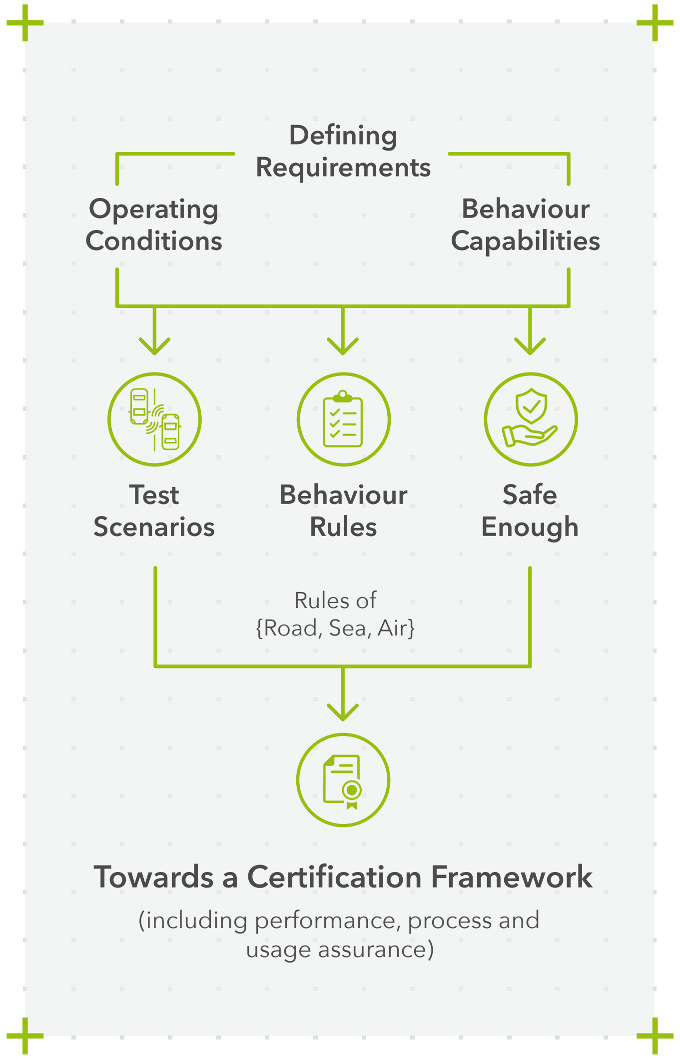
Therefore, performance assurance needs to be complemented by an evaluation of the process of development of the ATS.

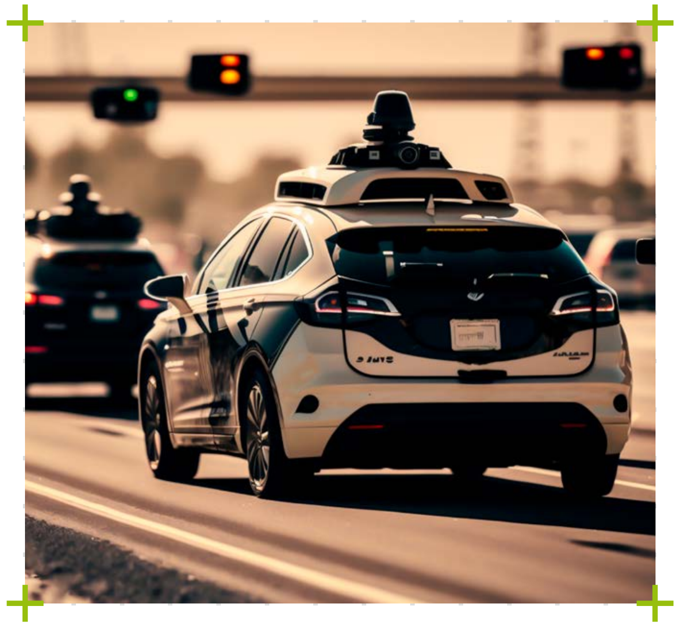
This is to give confidence that the ATS will be able to safely operate in scenarios that weren't part of the performance assurance process. This will be performed by an audit process.

Process assurance of the ATS involves the assurance of the process of development and post-deployment monitoring of the ATS. As a result, processes involved in each of the performance assurance pillars (test scenarios, test environment, and safety argument) and post deployment monitoring processes need to be assured via audit. This includes methodology for ODD definition, component and system requirements definition, formats or templates used for them and processes for risk acceptance definition.

One of the key findings through our working party discussions has been the need to consider commercial reality together with a robust assurance process. This is crucial in an emerging technology area like automated transport technology, as onerous processes may not be pragmatic to implement both from an industry perspective as well as regulators' perspective. At the same time, any introduction of ATS technology needs to meet a high safety threshold set by independent agencies.

The aviation industry currently has processes in place which enables it to share near-miss information, failure methods and off-board data recorders. These are shared between commercial stakeholders through regulator or independent agencies. Land and marine transport domains require similar approaches to be implemented. Other processes where similarities exist between transport domains include identification of a common failure mode taxonomy, decommissioning processes and a typical product development cycle (V-cycle of development).





4.3. Usage Assurance

The processes developed and illustrated as part of "process assurance" will be implemented by humans, both from the development of the ATS' perspective as well as the approval perspective (by regulators). Therefore, it is essential to ensure that those developing the ATS and approving the ATS have the appropriate skill set to perform the processes.

Process assurance and usage assurance form part of traditional safety management systems approach.

However, the skills required to audit both process and usage assurance are very different from existing systems' audits. It could be suggested that an audit of usage assurance is a check-box exercise with individual and organisation credentials being audited. However, an important aspect to highlight is that there doesn't exist any benchmark credentials for auditing assurance of ATS. The skills required for ATS assurance will be a departure from traditional skills of the workforce.

However, findings from the working party suggest that such skills are common across all transport domains and having a common pool of talent could be beneficial for the ATS ecosystem. It is important to create some type of certificate courses or qualification benchmark, similar to UKAS certification to provide confidence that the personnel performing the assurance activities have appropriate skillsets.

Understanding Assurance: Technical recommendations based on the working party discussions

- In order to have a scalable, consistent and manageable safety assurance framework, we argue that a safety assurance framework for ATS for each transport domain (land, air and marine) be underpinned by the Operational Design Domain (ODD) (i.e., operating condition) definition.
- For the effectiveness of the ODD-based safety assurance framework, we argue there is a need to have a minimum level of detail in the definition of the ODD of the ATS.
- With a diverse set of stakeholders in the ATS ecosystem for each transport domain, a common language for components of the safety assurance framework is needed. This includes a common language for ODD, test scenarios, safety metrics etc. While the actual keywords used for each transport domain may differ, the concepts for creation of the language could be similar.
- We argue the need to create a library of scenarios for each transport domain to enable sharing of scenarios across the ecosystem.
- We argue the need for a qualification scheme for virtual test environments (VTEs) which considers the qualification as a combination of virtual sensors and the virtual environment model. We argue that such a scheme would be similar across all transport domains, but the similarity thresholds for comparison between real-world and virtual-world outputs might differ for each transport domain and their corresponding use case.
- As uncertainties will always be present in every VTE, we argue the need for quantification of uncertainties to establish the performance parameters of the VTE.
- We argue for the creation of a codified definition for the Rules of the Road, Rules of the Air and Rules of the Sea, inspired by current definitions for human driven systems. The codification concept for each of the transport domains could be similar, based on an ODD and behaviour underpinning. Individual rule sets will be different in each of the transport domains.
- As demonstrated by use cases for ATS in the transport domain, remote operation (either monitoring, assistance, intervention or driving) will play a key role in the commercial deployment of ATS. We argue the need to establish general working principles for remote operation. These should especially focus on key challenges of human factors and connectivity for remote operation systems.

5. The Findings
Communicating Assurance

For the most part, people are not good at quantifying safety. We tend to focus on the outcomes of an unsafe event rather than its likelihood. That makes communicating the safety of any product or system challenging.

If we trust a regulatory body, based on their past performance, then we tend to believe their future predictions. In aviation, we trust that commercial aviation travel will be safe because the statistics are convincing and well communicated. We travel in aeroplanes made by a small number of manufacturers who, over decades have shown both that their products are safe, and that they rigorously investigate any accident and implement measures to stop the circumstances that caused the accident from reoccurring.

Similarly, we believe that travel in large commercially run ships or on national rail networks is safe because there are standards that both manufacturers and operators must adhere to. On the road, we trust that NCAP safety standards for the safety of car structures will minimise the consequences to the passengers of a crash if it does happen. For air, we trust the regulators (e.g., CAA or EASA or FAA), that any certified commercial aircraft is safe for humans to board on. For marine, we trust the class societies who qualify the vessels as safe in commercial deployment. It is important to highlight that here the trust in the ATS technology is institutional trust of the organisations approving the technology.

However, we have understandably come, as a society, to distrust safety claims provided by anyone who might benefit from our acceptance of their claims about safety. This means that convincing people that ATS are safe will start with the systems described above, but their validity must be made credible by support from regulators or other independent bodies.

Assuming that we have established the fact that an ATS is assured, one of the important questions to answer is that whether having the evidence that using the ATS was safe would make it more likely to be taken up or accepted by the users.

A sub-question to that is whether developers should explain the risks associated with their new ATS and its usage?

An interesting argument put forward in the working party meetings was that, with any new activity, people start off cautious, but familiarity leads to confidence and the risks are minimal and acceptable – until an incident happens, whereupon they might re-assess. Therefore, telling them the facts and their detail needs to be carefully thought through. The counter argument is that customers are becoming increasingly sophisticated. They are capable of, and more interested in, understanding the basis of how they might make their decisions.

Research has shown that users of automated technology need to develop an accurate understanding of the system's capabilities and limitations to enable them to use the technology in a safe manner. This has been termed as ***"informed safety"***[27]. It is the responsibility of the ATS developers to impart this knowledge to the users. This will include an accurate understanding of the ODD of the ATS by the user which suggests that the ODD needs to be defined in an accessible manner for the end user. A key motivation of communicating assurance is to build something trustworthy and communicating that as trust is one of the most important factors influencing the use of automation[28-32].

Having informed safety enables users to create an accurate mental model on how to safely use the automated technology.

Trust is defined as *"a history dependent attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability"*[27]. The reference to "history dependent" is particularly important for this work because prior knowledge about the system's capabilities and limitations affects an individual's attitude towards a system, thus affecting their trust.

Trust is said to be influenced by various factors including knowledge about the system, knowledge of certification, experience with the system, etc.[33].

One of the detrimental factors in the development of trust is over trust or mistrust in ATS. This occurs due to misleading or inaccurate communication of assurance. There are instances in industry where naming of non-automated or assistance systems can be misconstrued as an "automated" system e.g., Autopilot, Highway Pilot etc. Misrepresentation of the capabilities of automated systems has popularly been referred to as *"autonowashing"*[33]. Along with the ATS developers' responsibility to educate the user, we believe the government also has a role to check and promote consumer awareness. For the land domain, the joint paper between the Law Commission of England and Wales and the Scottish Law Commission on Automated Vehicles (now adopted by CCAV), has explicitly mentioned the misrepresentation of automated vehicle capabilities as a "criminal offence" (Recommendation 34)[34].

One of the working party findings was that to understand communicating assurance better, it is useful to consider three pillars of activities **[fig 8]**:

- **The Who:** Identifying and understanding the audience for communication.
- **The What:** Identifying content for communication.
- **The How:** Understanding the mechanism for communication.

Figure 8 Three aspects of communicating assurance

5.1. Communicating Assurance: The Who

The first pillar (who) focusses on identifying and more importantly understanding the audience for communication. Our findings suggest that there are a multitude of personas in different use cases who will need to be communicated to. This is consistent with the knowledge that there exists multiple stakeholders in the ATS ecosystem who will receive information about the safety assurance of the ATS. Identification of these personas and their characteristics is essential to establish the content and mechanism of communication (the other two pillars of communicating assurance).

The communication to the personas in the use cases will be done by government, non-government organisations (NGO), local authorities and industry. Potential government bodies will include regulators, certification authorities, and departments. Potential NGO bodies include special interest groups, universities, public sector research establishments, and standardisation bodies. Potential industry bodies will include ATS developers, ATS operators, Tier 1 suppliers and lobbying groups.

27 Khastgir, S., Birrell, S., Dhadyalla, G., & Jennings, P. (2018). Calibrating trust through knowledge: Introducing the concept of informed safety for automation in vehicles. Transportation research part C: emerging technologies, 96, 290-303.
28 Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics, 39(3), 429-460.
29 Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human factors, 39(2), 230-253.
30 Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human factors, 46(1), 50-80.
31 Walker, G. H., Stanton, N. A., & Salmon, P. (2016). Trust in vehicle technology. International journal of vehicle design, 70(2), 157-182.
32 Xu, J., Le, K., Deitermann, A., & Montague, E. (2014). How different types of users develop trust in technology: A qualitative analysis of the antecedents of active and passive user trust in a shared technology. Applied ergonomics, 45(6), 1495-1503.

33 Dixon, L. (2020). Autonowashing: The greenwashing of vehicle automation. Transportation research interdisciplinary perspectives, 5, 100113.
34 Law Commission of England and Wales, Scottish Law Commission, 2022. Automated Vehicles: joint report.

5.2. Communicating Assurance: The What and the How

The nature of interaction between the automated system and the various personas is different in different transport domains. The content and the mechanism of communication need to reflect the nature of this interaction. In order to understand the content and mechanism of communication, it is therefore essential to understand the nature of communication and human behaviour for the individual personas.

One of the classical frameworks to understand and influence human behaviour has been proposed by Jens Rasmussen which is still used widely in this field. Rasmussen's Skills-Rules-Knowledge (SRK) framework provides a guide and proposes abstraction hierarchies for human behaviour[35]. It refers to the degree of consciousness that an individual displays in undertaking a particular task. When users display knowledge-based behaviour (top-down) they are undertaking all activities in a completely conscious mode. This happens mostly when users are exposed to unfamiliar or novel situations. Rule-based behaviour is displayed when a conscious effort is involved in identifying the situation and matching it to a pre-coded or pre-defined rule for the user and the execution subsequently is done in a less conscious manner. Skill-based behaviour (bottom-up) refers to highly practiced physical actions with very little (if any) conscious monitoring or effort involved.

The introduction of automation increases system complexity, requiring users to demonstrate a top-down (knowledge-based) behaviour approach to accommodate for deviations in performance while receiving knowledge about the operational driving parameters (bottom-up knowledge). The significance of the abstraction hierarchies can be further illustrated by the fact that causes of failures or incorrect function are explained by a bottom-up approach whereas the reasons for the proper function are explained by a top-down approach[35].

This will enable to set expectations of the various personas who will form their own expectations based on the transport domain, use cases and situations.

To communicate assurance, trusted agencies need to understand the nuances of the SRK framework to create interventions in communicating at the appropriate levels for various personas. They will need to establish baselines and engineering thresholds which need to be communicated via mechanisms like kite marking[36] .

However, the communication from trusted agencies will help shape their expectations. In order to ensure accurate expectations, it is important that the personas are aware of the benefits, capabilities, limitations and risks of using the ATS.

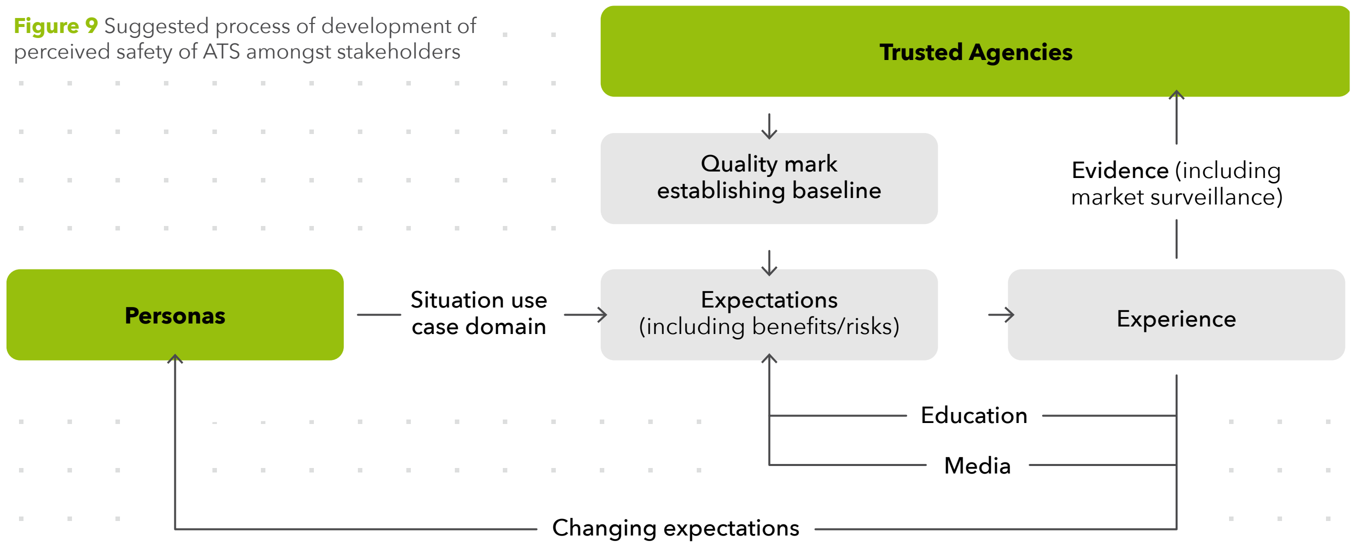


Figure 9 Suggested process of development of perceived safety of ATS amongst stakeholders

Based on their expectations, various personas will have their own experience of using the ATS. Experience (and market surveillance) with the ATS will provide further evidence for the trusted agencies to update the safety benchmarks (if needed) they have established and were communicating. At the same time, experience with ATS will also educate and make the personas more aware to enable them to calibrate their expectations to more appropriate levels **[fig 9]**.

It is important to appreciate that in some use cases of ATS technology, there doesn't exist a baseline to compare the technology against. For example, eVTOL (electric Vertical Take-Off and Landing) doesn't have a current baseline for safety in the public's eyes, however, self-driving vehicles have one as comparison to today's road safety levels. In such instances, the perception of safety becomes the truth irrespective of the accuracy of the perception or evidence to underpin it.

Therefore, proactive communication of assurance of ATS, is essential to ensure users have appropriate and accurate expectations from the technology. An important insight from the working party discussions was that the "independence of the agency communicating assurance" is key to enabling development of trust in ATS.

Another finding from the working party discussions was that we can't convert "perceived safety" thresholds into "engineering safety" thresholds. However, it was suggested that the principles of "perceived safety" need to be embedded in the process of establishing "engineering safety" thresholds. One way of doing this was consideration of off-nominal behaviours of ATS. This aligns with underpinning research which suggests to focus on communicating limitations of the ATS technology to develop trust in it.

It is important to appreciate the difference between imparting education, raising awareness and perceived safety.

Imparting education is an aspect of how to communicate assurance.

While the three concepts are related, there are subtle and importance differences between them.

Communicating Assurance: Technical recommendations based on the working party discussions

- We argue the need for the creation of a failure response mechanism for each transport domain (land, air and marine) to prevent distrust in ATS technology. Such a response mechanism should be guided by a common set of principles and should consider timely intervention (i.e., day 0 response, day 1 response, day 7 response etc.)
- We argue the need for communicating established remote operation principles in each transport domain with the ATS ecosystem stakeholders. Depending upon the audience, the content and mechanism of the communication will be different. However, the principles of creating the content should be similar across the transport domains.

35 Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE transactions on systems, man, and cybernetics, (3), 257-266

36 The BSI Kitemark™: https://www.bsigroup.com/en-GB/kitemark/

6. Understanding synergies and differences

Discussions in the working party meetings identified the synergies as well as differences between the transport domains (land, air and marine). The table below summarises these findings for '**understanding assurance**' and '**communicating assurance**'.

	Synergies across transport domains (land, air and marine)	Differences between transport domains (land, air and marine)
Understanding Assurance	Concept of Operational Design Domain (ODD) and behaviour and high level taxonomies. ODD definition methodology.	Detailed taxonomy elements of the ODD. Interaction between actors and the detailed behaviour taxonomy.
	Tools for managing requirements once linked with ODD and behaviour.	
	Process of codification of existing Rules of Road, Air and Sea into machine readable Rules for ATS.	Specific concrete rules for ATS will differ in each transport domain.
	Methodology for definition of "safe enough".	Safety thresholds for each transport domain will reflect the use case and appropriate risk appetite.
	Virtual Test Environment (VTE) qualification process.	Thresholds for acceptance (between the transport modes) of comparison between real-world outputs and virtual test environment outputs.
	Environment modelling (including climatology and sensor performance).	Each transport domain may have bespoke weather models to reflect the level of accuracy required for the domain and the corresponding use case.
	Data handling and exchange procedures between stakeholders. This should ensure both quality and protection of the data.	
	Standards for machine learning and artificial intelligence safety.	
	Security standards for physical and cyber security.	
	Remote operation guiding principles.	Remote operation standards for specific transport mode.
	Safety Management System procedures for the ATS and its independent assessment.	Specific regulatory controls for specific risks in each transport domain.
	Standards for organisation's qualification. Standards for individual person's qualification. Skills needed to implement and audit the safety management system.	
	Concept for network management (centralised or distribution) for the operation of the ATS.	
	Near miss definition for incidents and their subsequent management and learning process.	
	Failure or emergency response procedures (including crash investigation).	
Communicating Assurance	The concept behind the definition of automation levels in ATS.	Levels of User-in-Charge (UiC) and their expectations. This should reflect the difference in the nature of control between the user and the ATS in different transport domains. This will directly influence the operator training that may be required for each transport domain.
	Failure taxonomy for ATS.	
	Identification of personas to communicate safety. Principles of generating the information and corresponding communication mechanisms.	Concrete information communicated to each of the personas.

7. Areas of standardisation

The findings from the working party discussions made it evidently clear that there exist certain aspects of safety assurance (both ***understanding assurance*** as well as ***communicating*** assurance) which are common across the transport domains (land, air and marine).

However, the findings also illustrated that there exist areas which are specific to each of the transport domains (land, air and marine). In order to tackle these areas, we recommend the creation of a strategic cross-transport domain (land, air and marine) standards group to advise on the priorities for standards and help to initiate standardisation and pre-standardisation activities on the identified areas. This group could liaise with existing 'vertical groups' involved in standards activities relevant to automotive, aviation and maritime autonomy, share good practice across industry and promote synergies. Further benefits of such a group include the accelerated shared learnings and capabilities across regulators and industry and enabling the implementation of a cross-domain approach to a common assurance framework through standards and guidance, that enables technology growth and enhances the UK's global leadership in this sector.

One of the findings of the working party meetings was the urgency of outputs on common understanding of assurance for automated systems. There was also an appreciation that while standards are much needed, **it can take a substantial amount of time for the creation of standards.** There was widespread agreement on the **urgent need for guidance documents which explain aspects of assurances to industry and regulators. A guidance document (different from a standard) is advisory and not mandatory.** It tells the user how the regulator expects the user to show compliance and assure the ATS. The creation of such guidance documents needs to be undertaken by a group of practitioners which should include industry, academia and regulators.

In this case of standards activities, appropriate standards agencies (e.g., BSI, the British Standards Institution) should be empowered to undertake work on areas of standardisation and (in some cases) pre-standardisation activities to provide guidance to industry and regulators. Any such activity should be undertaken in consultation with existing standards users, industry, government, and relevant authorities.

In addition to the creation of standards, supplementary guidance is necessary on how the various standards apply in the assurance framework process.

As a result of the three working party discussions, and following on from the various technical recommendations, section 7.1 captures the various areas of standardisation. It is worth noting that some of the areas for standardisation are domain specific while others are domain agnostics.

7.1. Areas of standardisation: Understanding Assurance

Performance assurance

- Domain agnostic ODD-based assurance framework: leveraging existing concepts and processes in each of the transport domains and overlaying them with an ODD-based assurance concept. This could potentially lead to the creation of a quality marking scheme (e.g., BSI Kitemark™) for ATS.
- Domain specific ODD taxonomy: as each of the transport domains have physical characteristics and user interactions that are specific to the transport domain (land, air and marine), the ODD taxonomy will be specific to the domain.
- Language for defining ODD and defining test scenarios: to enable common understanding and ease of communication across stakeholders in the ATS ecosystem.
- Methodology for quantification of Virtual Test Environment (VTE) error quantification: this will enable the development of more trust in the test results using VTEs.

- Data output format for VTE: together with VTE error quantification, a common data output format for VTE will ease understanding of VTE results.
- Interfaces for VTE: ATS developers will use a variety of VTEs as part of their development. VTEs may potentially be modular.
- Domain agnostic approach: to define safe behaviour for automated transport systems.
- Domain specific rules for safe behaviour of automated transport systems (i.e., rules of the road, air and sea).
- Principles for remote operation (monitoring, assistance, driving) for ATS: it was suggested that most ATSs will at some part of their deployment require some level of remote operation (for initiation or intervention or regular monitoring). A common set of principles (especially incorporating the connectivity and human factors requirements) to ensure safety of such operations is needed.

Process Assurance

It was identified that each of the transport domains (land, air and marine) have a substantial process in place. However, benefits from harmonisation of the process could be derived. At the same time, it was identified that a lot of the processes (e.g., software development, systems engineering etc.) are already common across the transport domains. Specific areas of standardisation include:

- Response procedure to incidents: It was suggested that the response procedure (and its proactiveness) could potentially have an impact on the perception of safety of ATS. While some aspects of the response procedure to incidents may be transport domain specific, the wider majority were considered to be common across transport domains. Furthermore, it was suggested that this will require incorporation of timeliness and response / communication at regular intervals from the point of the incident.

Usage Assurance

- Personnel skill-set qualification scheme: ATS development and its safety assurance requires a variety of skillsets (e.g., systems engineering, AI, software development, system safety etc.). A skill-set qualification scheme will need to incorporate the diverse set of skills required and may lead to bespoke qualification for specific aspects of ATS development and safety assurance.

7.2. Areas of standardisation: Communicating Assurance

The working party discussion focussed on having standardised approaches to understand the content and the mechanisms of communication assurance to the diverse set of stakeholders.

- Clear and accurate **definitions** for levels of automation and levels of remote operation. While the precise classification may be domain specific, it is suggested that a domain agnostic approach to principles of defining levels of automation and remote operation would be beneficial for all stakeholders.
- Domain agnostic **failure mode taxonomy**: it was suggested that for ATS, it is important to communicate ATS limitations and failures to the stakeholders to help develop their accurate expectations of ATS. To this end, a common failure mode taxonomy across transport domains (land, air and marine) would help ease of understanding amongst stakeholders.
- **Quality marking scheme** (e.g., BSI Kitemark™) for automated transport systems: creation of such a scheme and the issuance of certification by a trusted and independent agency. While the creation of a quality marking scheme also benefits performance assurance, its communication about the independence of the issuing authority and the meaning of the "quality" will enable development of trust in ATS.

8. Our Key Recommendations

After over 200 person days of discussions and 10 working party meetings with participants from over 30 national and international organisations, the following are our key recommendations to enable the creation of a cross-domain approach to safety of automated transport systems across land, air and marine.

Establishing Safety Levels

- **Develop a safety assurance framework for automated systems in land, air and marine.**

  This should be based on the concept of Operational Design Domain (ODD) (i.e., the limit of operating conditions) which allows for a manageable, consistent and scalable assurance process within each domain, and interoperability between domains.

- **Create a qualification process for Virtual Test Environments (VTE) to enable trust in evidence from virtual testing.**

  The use of Virtual Test Environments (VTE) will be essential for safety assurance of ATS (in each domain). Such a qualification process needs to ensure it confirms qualification of all the virtualised components of the test environment.

- **Create standardised taxonomies and definitions for each component involved in the assurance process.**

  The wide variety of stakeholders involved in the safety assurance process demands a common understanding of the various components involved in the assurance process (i.e., ODD, test scenarios, safety metrics and thresholds). This requires clear commonalities where possible (e.g., levels of automation etc.).

- **Coordinate various programmes and government initiatives across all domains. There is a need to prove and consolidate the approach for the safety of automated transport systems.**

Communicating Safety

- **Create a common set of framework principles to communicate safety to all stakeholders in each transport domain.**

  The message will need to be tailored to the relevant audience and the content of the communication will need to vary depending on the type of stakeholder (e.g., public, developers, regulators, insurers).

- **Encourage independent organisations to take a proactive role in communicating safety of ATS.**

  The credibility and independence of the organisations and people communicating safety is key. Establishing trust and credibility in the messaging through reputable organisations and individuals will help to grow trust in the technology.

Contributors to the working party discussions

1. Aerospace Technology Institute (ATI)
2. Aurrigo Driverless Technology
3. Blue Bear Systems Research
4. BSI (British Standards Institution)
5. Centre for Connected and Autonomous Vehicles (CCAV)
6. Civil Aviation Authority (CAA)
7. Collins Aerospace
8. Connected Places Catapult
9. Department for Transport, HM Government
10. Driver and Vehicle Standards Agency (DVSA)
11. European Commission Joint Research Centre (JRC)
12. Imperium Drive
13. King's College London
14. Law Commission of England and Wales
15. Lloyd's Register
16. Maritime and Coastguard Agency (MCA)
17. Mathworks UK
18. Met Office
19. Military Aviation Authority
20. National Physical Laboratory
21. National Highways
22. NATS
23. Oxbotica
24. Port of Dover
25. Reed Mobility
26. SysElek
27. Tata Consultancy Services
28. Trilvee
29. UKRI Future Flight Challenge
30. UKRI Trustworthy Autonomous Systems (TAS) Hub
31. University College London
32. University of Leeds
33. University of Leicester
34. University of Plymouth
35. Vay
36. Vehicle Certification Agency (VCA)
37. Wayve
38. WMG, University of Warwick


Remote Driving
Challenges and opportunities across transport domains


WMG
THE UNIVERSITY OF WARWICK
CATAPULT
Panel Discussion 1:
Regulation of remote driving across domains, challenges and considerations
Connor Champ, Lawyer, Law Commission
Sarah Smith, Lawyer Law Commission
David Wong Senior Technology and Innovation Manager, Society of Motor Manufacturers and Traders
Chair:
Steve Kendall, Product Manager for Specialist Vehicle Examinations, Driver & Vehicle Standards Agency
Konstantinos Mattas, Research Fellow European Commission - Joint Research Centre
Katrina Kemp, Maritime Autonomy Policy Lead Maritime and Coastguard Agency
Dr Andrew Berry, Emerging Policy Specialist, Civil Aviation Authority

For more information about this report, Verification & Validation team or to work with us, please contact:

Professor Siddartha Khastgir
S.Khastgir.1@warwick.ac.uk
SafeAutonomy@warwick.ac.uk

WMG
THE UNIVERSITY OF WARWICK

warwick.ac.uk/wmg

CATAPULT High Value Manufacturing

UK Research and Innovation