THE | AUTONOMOUS

Chapter Event Safety & Architecture

co-hosted by TTTechAuto

EXECUTIVE SUMMARY

On April 2nd, 2020, The Autonomous together with TTTech Auto hosted a virtual Chapter Event on "Safety & Architecture". The event featured six presentations and two panel discussions. A moderator managed the interaction between the audience and the speakers. Indeed, the audience submitted numerous questions that were answered by the presenters or as part of a post-event survey. The event focused on two main topics: (i) system architectures for safe automated driving (AD) and (ii) safe trajectories for AD. This report summarizes the presentations, panel discussions, the Q&A, and the results of the post-event survey.

Focus I: Safe AD Architectures

The topic aimed to answer the following question: "How can we establish safe architectures for AD?" Three high-quality keynotes from industry and academia were presented on this matter, and 16 technical questions were thoroughly discussed - some of which are:

- Is AI a mandatory component in Fail-Operational/Fail-Degraded (FO/FD) AD Architectures?
- How do you ensure the overall AD system's safety in case of frequent over-the-air updates?
- What are the pros and cons of the most common FO/FD architectures?

Furthermore, a post-event survey resulted in the following data:

- Participants considered costs, proving safety, and real-time capabilities as the main challenges in the development of future FO/FD AD architectures.
- When it comes to the scalability of FO/FD AD architectures, 51% think they should not be scalable, whereas 41% think they should, i.e., from SAE L2 AD to SAE L5 AD.
- 55% considered Doer/Checker with fallback the most suitable FO/FD architecture for AD, whereas 17% think Triple Modular Redundancy is more suitable, and 14% suggest others.
- A majority (59%) consider that there is no need for completely independent sensor HW for each redundant channel in FO/FD AD architectures. Whereas 38% think there is a need.

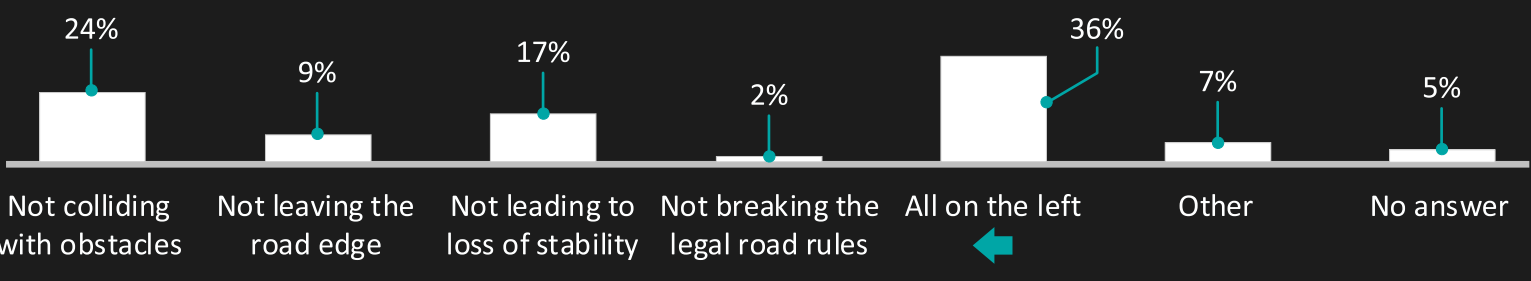
Focus II: Safe AD Trajectories

"What are the criteria that an AD trajectory should meet to ensure that it is safe?" was the main question addressed in this topic. Once again, three high-quality keynotes were presented, and 14 technical questions were passionately discussed – a portion of which are:

- How can a safety monitor (i.e., the Checker in Doer/Checker) qualify a trajectory as safe?
- Will the safety monitoring functionality be something that is provided by Tier-1 suppliers as "Plug-and-Configure" functionality, or will it stay OEM responsibility?
- Which standards and initiatives focus on safe behavior specifications?

The post-event survey also focused on these criteria and resulted in the following data:

- With regards to whether common/standardized interface definition for AD trajectories is needed, 73% have answered "Yes", 17% "No", and 10% have provided no answer.
- The chart below depicts the results for: "which criteria an AD trajectory should meet to ensure it is safe?"



BACKGROUND AND EVENT DETAILS

The Initiative

For all actors involved in the development of autonomous mobility solutions, who position safety as a fundamental value of their products - The Autonomous is a knowledge ecosystem - that generates new knowledge and technological solutions to tackle key safety challenges that shape the future of safe autonomous mobility. Complementary to standardization organizations that establish uniform engineering or technical criteria, methods, and processes, The Autonomous will develop Global Reference Solutions for autonomous mobility that conform to relevant standards and facilitate the adoption of these solutions on a grand scale. The benefits The Autonomous will provide to the partners of the ecosystem are:

- Development of safe and best-in-class AD solutions thanks to the wisdom of the crowd;
- Reduction of potential liability risk by (i) tightly working with government and regulatory institutions and (ii) fostering user-centric solutions that enable incremental development of autonomy;
- Reduction of development costs by (i) developing modular and reusable Global Reference Solutions and (ii) sharing the development efforts;
- Accelerating the learning curve by collectively learning from individual failures and field observations;
- Joint definition of state-of-the-art and state-of-practice.

Towards this vision, in 2020, The Autonomous is hosting a series of workshops - "The Autonomous Chapter Events" - to facilitate discussions among experts and take the first steps towards the targeted Global Reference Solutions. The first Chapter Event titled "Safety & Architecture" was hosted by The Autonomous, together with TTTech Auto.

Event Details

Presentations

Focus I: Safe AD Architectures

- Architecting AD Systems with AI | Riccardo Mariani | NVIDIA
- System and Software Architecture for AD | Simon Fürst | BMW
- Architecting, Verifying, and Validating AD | Martin Törngren | KTH Royal Institute of Technology

Focus II: Safe AD Trajectories

- Models, Metrics, and Assumptions in Safety Assurance | Jack Weast | Intel
- Requirements for Safe Trajectories | Wilfried Steiner | TTTech Auto
- AV Trajectories: Newtonian Mechanics vs. The Real World | Philip Koopman | Edge Case Research

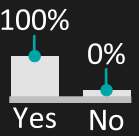
Event Statistics

Facts

- 258 event registrations
- 115 companies & Institutions attended
- YouTube livestream:
  - 501 unique views
  - 120-179 concurrent viewers
- 30 technical questions addressed

Feedback

- 50 participants provided feedback
- Did the event meet your expectations?



- How would you rate the event? ★★★★★½

Technical Report

Chapter Event Safety & Architecture

Edited by
Ayhan Mehmed

The Autonomous
May 2020

Contents

1 | The Initiative

As autonomous mobility is moving closer to becoming a reality, safety and trust concerns prove to be the main hurdle in the way of reaching broad acceptance. OEMs and technology suppliers (Tier 1, 2 & 3, and others) cannot overcome the safety challenge and the necessary investment costs with a "go it alone" approach. Therefore, the autonomous mobility industry and other relevant institutions need to come together and show significant efforts in prioritizing and ensuring safety on all technological levels, as well as set common technical and legal standards. Towards this, TTTech Auto initiated The Autonomous - an open platform that brings together actors from the autonomous mobility ecosystem to align on relevant safety subjects.

1.1 Vision

> *Create a safer, more livable,*
> *and more sustainable future.*
>
> — The Autonomous

For all actors involved in the development of autonomous mobility solutions, who position safety as a fundamental value of their products - **The Autonomous is a knowledge ecosystem** - that generates new knowledge and technological solutions to **tackle key safety challenges** that shape the future of safe autonomous mobility. Complementary to standardization organizations that establish uniform engineering or technical criteria, methods, and processes, The Autonomous will develop **Global Reference Solutions** for autonomous mobility that conform to relevant standards and facilitate the adoption of these solutions on a grand scale. The benefits The Autonomous will provide to the partners of the ecosystem are:

- Developing safe and best-in-class solutions for Automated Driving (AD) challenges thanks to the wisdom of the crowd;
- Reduction of potential liability risk by (i) tightly working with government and regulatory institutions and (ii) fostering user-centric solutions that enable incremental development of autonomy;
- Reduction of development costs by (i) developing modular and reusable Global Reference Solutions and (ii) sharing the development efforts;
- Accelerating the learning curve by collectively learning from individual failures and field observations;
- Joint definition of state-of-the-art and state-of-practice.

Furthermore, the work products of The Autonomous are expected to serve as further input to existing standardization activities and may also result in new standardization projects.

1.2 Mission

Towards the above-defined vision statement, The Autonomous will:

- Provide a diverse and balanced knowledge ecosystem for autonomous mobility;
- Set the stage for open discussions on main technical and architectural questions where controversial approaches can be freely discussed;
- Act as an interface between industry requirements, standardization, regulation bodies, and academic research in safe autonomous mobility. Collectively identify important gaps in the field and focus the efforts;
- Build consensus on major safety solutions within the automotive industry;
- Generate high-quality know-how and Global Reference Solutions compliant to relevant standards in autonomous mobility;
- Facilitate the adoption of the Global Reference Solutions on a grand scale by placing them into relevant standards as solutions compliant to their requirements.

1.3 Approach

Current Approach

The development approach of automotive systems has remained unchanged over many years. Generally speaking, a car manufacturer (OEM) and its suppliers (Tier 1, 2 & 3, and others) cooperate and then compete with other manufacturers in providing better solutions and products (see Figure 1). This approach has worked well for developing standard, well constrained, and deterministic automotive embedded systems like Anti-Lock Braking System (ABS), Engine Control Units (ECU), and others.

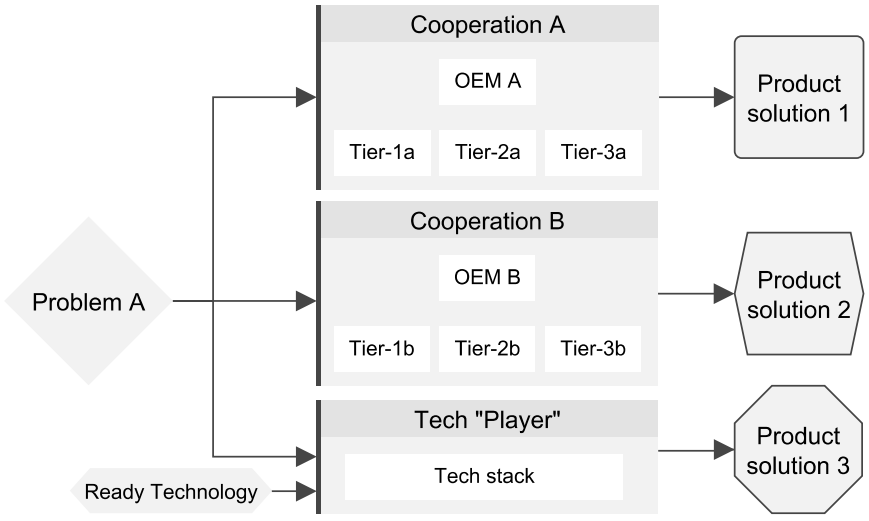


Figure 1: Current development approach of automotive systems.

However, the approach is sub-optimal when it comes to the development of upcoming SAE Level 3 - Level 5 Automated Driving Systems (ADS). The rationale for this is (i) the novelty and high complexity of the AD systems, (ii) unprecedented high development costs, and (iii) different technical solutions will likely not align in a common state of the art.

Proposed Approach

To reduce the development cost, a shift from many interdependent cooperation groups (where cooperation groups compete with each other on providing a better solution for a given problem) to a single, larger, and more diverse knowledge ecosystem where partners collaborate towards a single shared goal is necessary (see Figure 2). Such an approach will enable (i) the development of safe and best-in-class products, (ii) an ecological and sustainable development, and (iii) faster development autonomy. Furthermore, in addition to car manufacturers and technology suppliers, The Autonomous also invites stakeholders from governmental, academic, regulatory, and standardization institutions in order to ensure an integrated view.

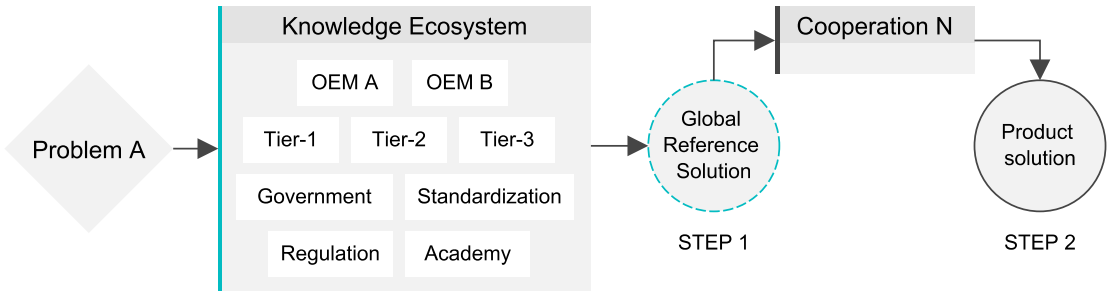


Figure 2: Proposed approach for development of future AD systems.

In “STEP 1” of the proposed approach, the partners of the knowledge ecosystem will work together on Global Reference Solutions that conform to relevant standards. The notion of the Global Reference Solutions is to cover all relevant problems in the development of future AD systems. Hence, more than one reference solution will be available, i.e., ranging from Fail-Operational/Fail-Degraded (FO/FD) architectures to verification and validation (V&V), runtime verification approaches, sensor and sensor fusion configuration, and others. In “STEP 2” of the proposed approach, the partners of the ecosystem will be able to individualize the Global Reference Solution to their needs and therefore keep the competition “alive".

1.4 Roadmap

In 2020, The Autonomous is organizing a series of *virtual* technical workshops, also known as “The Autonomous Chapter Events”, to facilitate discussions among experts and work towards the target Global Reference Solutions. Figure 3 presents a summary of the Chapter Events planned for 2020. While the scope of the Chapter Events will be

further broadened by adding other relevant topics, the list below summarizes the current status:

- Chapter Event Safety & Architecture: 2nd of April, 2020 with co-host TTTech Auto;
- Chapter Event Safety & Artificial Intelligence (AI): 5th of June, 2020 with co-host Five;
- Chapter Event Safety & Security: 22nd of June, 2020 with co-hosts Infineon, Secunet, and Integrity Security Solutions;
- Chapter Event Safety & Regulation: 9th of July, 2020 co-hosted with Posser Spieth Wolfers & Partners (PSWP).

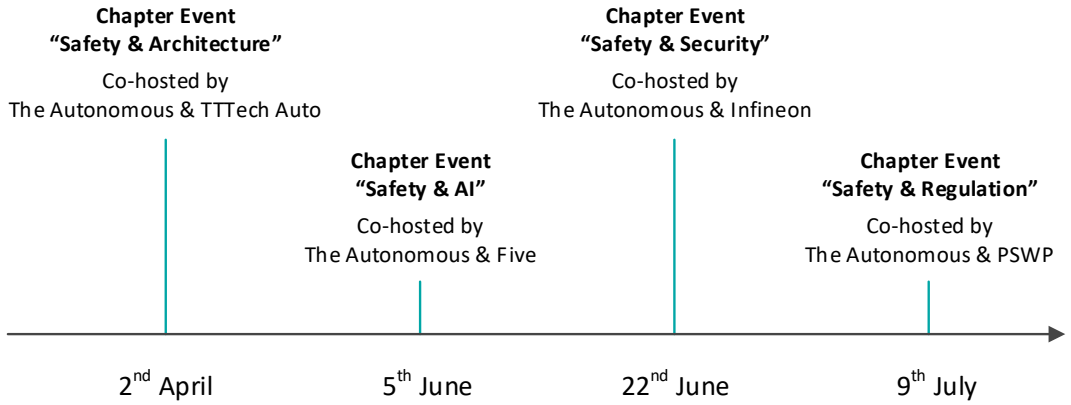


Figure 3: Summary of planned Chapter Events for 2020

The target outcome of each Chapter Event is a high-quality content summarized in reports. The current report is a summary of Chapter Event Safety & Architecture.

2 | Chapter Event Safety & Architecture

2.1 Scope and Topics

The Safety & Architecture Chapter Event is focusing on the following two topics:

- **Focus I: Safe AD Architectures**
  Conduct a concretization work on safe architectures for AD. The main question to be answered is "How can we define safe architectures for AD?" and as a starting point, we focus on the following topics:
  - Current architectural approaches in ADAS[1] and ADS,
  - Architecting FO/FD behavior,
  - Decomposition examples of the proposed FO/FD architectures to achieve the necessary Automotive Safety Integrity Level (ASIL).
- **Focus II: Requirements for safe trajectories**
  - Definition of common/standardized interface of a trajectory,
  - What are the criteria that a trajectory should meet to ensure that it is safe,
  - Review of current approaches for verifying the safety of trajectories (Responsibility Sensitive Safety, Safety Force Field, and others).

2.2 Event Statistics

Figure 4 summarizes the facts about the event and the feedback given by the participants. In particular, 258 registrations were made for the virtual event. The participants were from 115 different companies/institutions. The YouTube livestream had in total 875 views, out of which 501 were unique views. Throughout the four-hours event, there were 120 to 179 concurrent viewers. Last but not least, 109 questions were asked by the audience, of which 30 were addressed (see Section 3 and Section 4 for the summary of answers). Fifty participants provided feedback after the event, where 100% of them said "yes" when they were asked whether the event met their expectations. The participants also rated the event with five-and-a-half stars out of six.

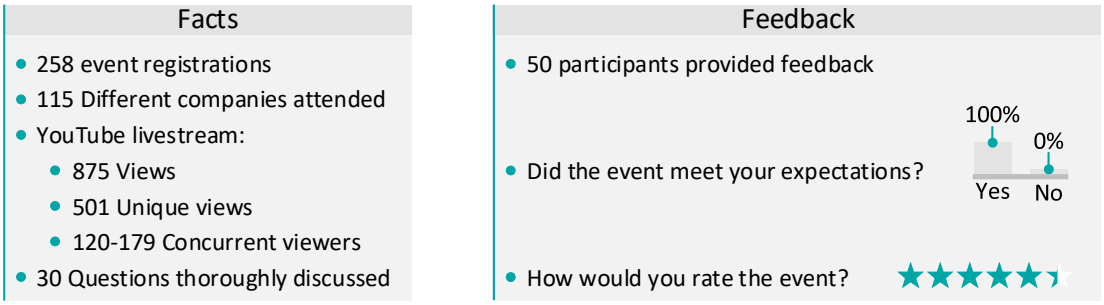


Figure 4: Facts about the event and feedback from participants.

[1] Advanced Driver Assistance Systems (ADAS).

3 | Focus I: Safe AD Architectures

3.1 Talk 1: Architecting AD Systems with AI

Riccardo Mariani
Vice President of Industry Safety, NVIDIA

Summary

In order to realize widescale deployment, AI-enabled AD systems demand software-defined end-to-end architectures that are functionally safe. This presentation delves into some of the key elements and challenges associated with AD systems, including functional safety design and validation. It will also look at today's AD systems standardization landscape and ongoing industry collaboration activities.

Addressed questions

Q1 Is AI a mandatory component of future FO/FD AD architectures? How do you guarantee they will not jeopardize the safety of the whole system?

Answer by Riccardo Mariani
AI is a key element of future AD systems as they complement classical algorithms (e.g., Kalman filter) at the event of complex driving scenarios. Indeed, ensuring an AI system (i.e., algorithm, software, and hardware) is safe, is a fundamental challenge. Currently, there are three complementary approaches to tackle this challenge. First, an AI system has to be designed according to functional safety requirements, ISO 26262 (FuSa) [1]. This implies the process with which one develops the AI algorithm and the tools with which one trains the AI algorithms. Standards as ISO/PAS 21448 (SOTIF) [2], ISO/CD TR 4804 (SaFAD) [3] have an Annex on designing an AI system in a safe way. Second is the appropriate V&V of the AI algorithm. The third approach is to use monitoring to verify the safety of the AI-based AD systems at runtime (i.e., when the vehicle is operating on public roads).

Q2 How do you ensure the overall AD system's safety in case of regular over-the-air updates. Will a re-validation be necessary?

Answer by Riccardo Mariani
Indeed, future AD systems are foreseen to have shorter update cycles, and the question of "how to ensure AD system's safety after an update?" is receiving significant attention in the industry, academia, and standardization. It is expected that most of the update cycles will include feature updates (e.g., a better lane detection algorithm, extending the Operational Design Domain (ODD) by enabling the functionality at different weather conditions, etc.) that are well segregated from the safety-critical part. In the case where the over-the-air update affects the safety-critical part, an upfront V&V of the system is necessary before releasing the update. Hence a well-defined and effective end-to-end V&V process, as well as a capable and efficient simulation environment, is essential.

3.2 Talk 2: System and Software Architecture for AD

Simon Fürst
Principal Expert Automated Driving Technologies, BMW Group

Summary

Currently, numerous OEMs are working on SAE Level 3 Highway Pilot Systems to be released in the next vehicle generations. This new customer functionality will become the high-end feature of ADAS, ranging from SAE Level 1-3. Due to this fact, a scalable approach for a system, ECU, sensor, and software architecture is required so that features of higher SAE Level can be added on top while re-using already existing features. This requires, on the one hand, an extensible sensor setup with additional sensors complementing the entry-level sensor setup. On the other hand, scalable SoC architectures must be used to enable the re-use of software from entry-level SoCs up to high-end SoCs. However, this scalable approach places high demands on safety. As of today, all automotive safety architectures are fail-safe, i.e., whenever a failure occurs, the system goes immediately into an electrical shutdown, and typically a mechanical fallback is in place. In SAE Level 3 ADS, the human driver hands over responsibility to a system that no longer requires a human fallback to be in place at every second. As a direct consequence, new safety architectures are needed. They require a fail-degraded capability so that in case of a single failure, the system can still actively hand-over to the human driver or execute a minimum risk maneuver to bring the vehicle into a safe state. To fulfill such requirements, current architectures of AD systems have a primary and a secondary safety channel as well as a fallback channel. The primary channel calculates the "comfort" trajectory that gets executed in regular operation. The secondary channel has the main task of evaluating if the trajectory of the primary channel is free of any collisions. Both channels use the same sensor setup and continuously cross-validate each other to monitor if each channel is working correctly. To further enhance redundancy, sensor fusion takes place on different levels in these two channels. The fusion levels are ranging from raw data sensor fusion up to fusion on the level of object lists. In case the primary and secondary channel have a major misfit of their trajectories, the fallback channel gets activated. This channel uses a minimum sensor setup and a safety-trajectory originating from the last failure free cycle to bring the vehicle into a safe state. Additionally, the primary/ secondary channels are powered by the main power supply, while a redundant power supply powers the fallback channel. This safety approach realizes multiple levels of redundancy and supports the detection of random hardware faults, sporadic and systematic software faults, as well as SOTIF issues. Further details on BMW's approach on safety for SAE Level 3 ADS can be found in the BMW Group Safety Assessment Report, see [4].

Addressed questions

Q3 In a FO/FD architecture that consists of a Doer (i.e., Primary channel), Checker (i.e., Secondary channel), and Fallback components, how does the Checker verify the correctness of the Doer?

> **Answer by Simon Fürst**
> To begin with, the Doer and the Checker have to be highly independent. The verification of the Doer is done on multiple stages, and not only at the end. In the most common implementation, the Doer generates a trajectory that aims at ensuring a safe and comfortable drive. The task of the Checker is to verify whether the generated trajectory will not lead to an unsafe operation (e.g., collision to other obstacles on the road).

Q4 How does the supervision by Checker work when the performance of the machine learning-based Primary ASIL channel (Doer) outperforms the Checker?

> **Answer by Simon Fürst**
> The Checker is mainly designed to be safe and perform verification tasks of the trajectories generated by the Doer. Therefore it has lower precision than the Doer when it comes to certain capabilities. For instance, concerning localizing objects on the road, the Doer can locate an object with a precision of few centimeters. When it comes to the Checker, its precision can be lower, while its reliability is higher. Under all circumstances, the Doer's trajectory must be verified if it is safe by the Checker. Hence, the Doer needs to cross-check whether the trajectory is not driving at a non-reasonable distance to other objects on the road.

Q5 Which quantification criteria are used to justify ASIL B for the AI-based components, i.e., how to argue ASIL B for AI-based components of an AD system?

> **Answer by Simon Fürst**
> When it comes to a Highway Pilot AD feature, not too many components of the AD system are AI-based. AI is mostly used in image processing and in a limited way in the planning part of the AD functionality. The upcoming release of ISO/CD TR 4804 (SaFAD), Annex B, will provide detailed information on the process of justifying a certain ASIL for AI-based components. Furthermore, it is worth noting that the question of arguing the safety of AI-components is not a completely solved problem. There are research and standardization activities that work on a common understanding of the mentioned problem.

3.3 Talk 3: Architecting, Verifying and Validating AD: Observations and Open Questions

Martin Törngren
Prof. at KTH Royal Institute of Technology

Summary

The Mechatronics division at KTH in Stockholm addresses challenges posed by future Cyber-Physical Systems (CPS) and specifically with their incarnation in terms of various types of automated machines, including highly Automated Vehicles (AVs). Key research questions addressed in the division are:

- What makes a compelling, comprehensible, and valid safety case for a highly AV?
- What makes an appropriate reference architecture for AVs that could be instantiated for different types of AVs, providing capabilities to ensure cost-efficient risk reduction and the right balance between availability/safety?
- What represents an appropriate methodology for model-based evaluation of highly AV?

With these questions as a starting point, this brief talk will:

- Provide observations regarding the evolution of the area of AVs and the need for collaboration.
- Elaborate the need for, and challenges in, system formalization to support verification, with recent findings from a survey of methods to derive critical scenarios.
- Discuss safety supervisor architectures, considering single vs. collaborative architectures and safety concepts.
- Reference architectures to ensure cost-efficient risk reduction and the right balance between availability/safety for different extra-functional requirements.

Addressed questions

Q6 Tools for system modeling and simulation tools are quite mature. Is there not a huge gap in tools for modeling and analyzing ODDs, safety cases? What do you use?

Answer by Martin Törngren

Yes, tools for system modeling and simulation are indeed mature. However, AD pushes the boundaries in at least two ways:

1. The resulting systems integrate technologies and systems at an unprecedented level. This means that while individual domain/discipline models/tools may be mature - their combination may not be.
2. Certain areas are newer and expanding, for example, in terms of needs to describe environments (and the ODD) and to model more complex safety cases.

Q7 What are the most important characteristics of a "trustworthy" CPS?

> **Answer by Martin Törngren**
> Trustworthiness is a well-established term. It has to do with how we perceive a system. Just like dependability, it has multiple properties. We usually include at least the following when we refer to trustworthiness: safety, security, reliability/availability, predictability, and privacy.

Q8 Which process did you follow for building the safety case for the research concept vehicle?

> **Answer by Martin Törngren**
> We are currently conducting the safety case for our research concept vehicle with the transport authorities in Sweden. We are applying established engineering practices for system safety; we will be publishing our safety case once this is completed.

3.4 Panel Discussion on Safe AD Architectures

Addressed questions

Q9 An essential part of a FO/FD architecture is the so-called Selector, which is responsible for switching between the Doer and the Fallback sub-systems of the FO/FD AD system depending on the verification results from the Checker. In such a setup, the Doer and the Fallback are developed to ASIL B, whereas the Selector is developed to be ASIL D. Why is that?

> **Answer by Simon Fürst**
> According to ISO 26262 Part 9, the ASIL of each safety requirement and respectively, the ASIL of the ADS' item can be lowered by decomposing it into two redundant requirements. For example, an ASIL (D) ADS can be developed by implementing two ASIL B(D) AD sub-systems - i.e., Doer ASIL B(D) and Fallback ASIL B(D). The Selector needs to be ASIL D(D) as it an item that connects the Doer ASIL B(D) and Fallback ASIL B(D). See more decomposition examples in ISO 26262 Part 9.

Q10 How to achieve an ASIL D in the Selector?

> **Answer by Simon Fürst**
> The Selector functionality is implemented in a well-known inherent system - that is, the system responsible for the Electronic Stability Control (ESP) and Anti-Lock Braking (ABS). The hardware and software of this system are developed to ASIL D, according to ISO 26262.

Q11 Which standards and initiatives focus on safe behavior specifications?

Answer by Riccardo Marianni
When it comes to safe behavior specification, different technology providers have differing opinions. This is expected, as there is no single rule that, if followed, will guarantee the safe behavior of the system. Hence, it is also important to focus on a methodology for defining the rules/specifications for safe behavior. The currently developed IEEE P2846 standard on "Formal Model for Safety Considerations in Automated Vehicle Decision Making" is focusing on defining such methodology [5]. Moreover, standards as ISO/PAS 21448 (SOTIF), ISO/CD TR 4804 (SaFAD) are also focusing on safe behavior specification.

Q12 Is formal verification an option to be used in the Checker?

Answer by the Editor
Formal verification is a well-known and widely used approach for verifying the safety of a system during the development phase of a system. The Checker, on the other side, implements runtime verification approaches that are applied during the system's deployment phase, e.g., by continuously checking the current execution of a system. The verification approach executed in the Checker has (i) to run in real-time (e.g., execution cycle in the range of milliseconds), as well as (ii) to not add a lot to the overall AD system's end-to-end latency (e.g., from sensing to actuating). One of the challenges in implementing a formal verification approach in the Checker is the longer time for executing the formal verification (typically in the range of seconds).

Q13 Can formal verification be used to verify the safety of the Checker?

Answer by Riccardo Mariani
Formal verification is one way for verifying the safety of the Checkers that are going to implement given safety rules. IEEE P2851 is also addressing this topic [6].

Q14 Regarding testing software and AI solutions, how can one ponder the efficiency of the simulations compared to real scenarios?

Answer by Riccardo Mariani, Simon Fürst, Martin Törngren and refined by Editor

There is no metric for measuring which approach is better: e.g., the simulation or the brute-force testing on a public road. The two approaches are complementary and have their benefits and drawbacks. The current state of practice applies simulation in the early stages of AD system's development as it allows:

- Testing the system on a diverse number of scenarios. The catalog of scenarios can be stored and extended with each new experience.[a]
- Thanks to powerful computing centers, the simulations can run "x" times faster than real-time. Thus, allowing rapid testing of a vast number of miles in a shorter time in comparison to public road testing.
- Fast re-testing after new developments in hardware and software.

The simulation approach brings a complexity challenge that grows with the increase of the quality of the simulated environment: for that reason, it requires very high-performance computing (e.g., GPU clusters). One also has to take care not to over-trust the simulation models and their results. Therefore it is essential to validate the models and the simulation tools.
On the other hand, testing on public roads (or specialized for testing closed roads) compensates for the lower fidelity of the simulations, i.e., simulation tools are quite good nowadays. However, they cannot reproduce the real environment with 100% accuracy. In addition, testing on public roads helps with discovering unknown scenarios that have not been foreseen before. A major part of the final testing of the AD system is done on simulation, but final approval and release testing will always be on a closed testing ground or public road.

[a] Real-world driving datasets can be one way to reduce the gap between "pure" simulation and public testing. Paper [7] provides an overview of 37 publicly available datasets.

Q15 How can one assure the correct detection of anomalies in the AD system?

Answer by Riccardo Mariani, Simon Fürst, Martin Törngren and refined by Editor

Anomalies from the functional safety perspective can be mitigated by developing the AD system according to the well-proven ISO 26262 process. When it comes to anomalies of the intended functionality of the system, the ISO/PAS 21448 (SOTIF) standard is the right place to look into.
Furthermore, in the context of AI, there is ongoing research in developing approaches that highlight when anomalies of the AI algorithm occur. Another research activity is to use AI for detecting anomalies. Finally, cyber-security should not be forgotten. A properly executed cyber-security attack may be able to inject a fault that leads to abnormal functioning of the system. Therefore the AD system should be well protected against internal and external attacks.

Q16 What are the pros and cons of the most popular FO/FD architectures - e.g., dual standby fault-tolerant and triple modular redundancy (TMR) architectures?

Answer by Riccardo Mariani, Simon Fürst and refined by Editor
Each of the architectures has its advantages and disadvantages. Disruptors companies are in favor of the classical TMR architecture that is well known from the avionics. The majority of more traditional technology providers and OEMs are in favor of the dual standby fault-tolerant architecture (e.g., Doer/Checker with Fallback). At present, there is no one single reference solution for a FO/FD AD system architecture.
Moreover, we need to understand that the existing FO/FD architectures are only focusing on functional safety. However, it is well-known that the key challenge of achieving the AD functionality is in the safety of the intended function[a]. Therefore the existing architectures need to be adapted. This includes the use of different types of algorithms: from classical ones to newly developed AI algorithms. Diverse sensor setups for each of the channels is also a matter to be addressed.

[a] See [8] for more information about the difference between Functional Safety (FuSa)(i.e., ISO 26262) and Safety of the Intended Functionality (SOTIF)(i.e., ISO/PAS 21448).

4 | Focus II: Safe AD Trajectories

4.1 Talk 4: Models, Metrics and Assumptions in Safety Assurance

Jack Weast
Sr. Principal Engineer, Intel

Summary

In this talk, Jack Weast provides an update on the latest industry efforts to define common metrics for assessing the safety of AV using models like Responsibility Sensitive Safety, and a unique perspective on the critically important role that simple, human-understandable assumptions have in safety assurance claims that are understandable by the public and government alike.

Addressed questions

Q17 Will the safety monitoring (i.e., Checker) functionality be something that is provided by Tier 1 suppliers as "Plug-and-Configure" functionality, or will it stay OEM responsibility?

> **Answer by Jack Weast**
> Safety is an attribute that has to be considered at the beginning of the system development - i.e., at the conceptual and design phase. Thus the safety monitoring functionality should not be an add-on component that is integrated at the very end of the system development. Instead, it should be an integral part starting from the very beginning of the development.

Q18 If the majority of accidents are caused by humans, then why do we try to make the vehicles as good as humans? Why do we not aim for better?

> **Answer by Jack Weast**
> When it comes to comparing the overall number of accidents made by a human driver and AV, the latter should be better (i.e., cause fewer accidents). To begin with, shifting the driving functionality from the driver to a system already brings benefits: e.g., a system will not be distracted or drive drunk. However, if a system cannot drive in a human-like manner, it is likely to contribute to more accidents than it could be preventing. Therefore it is necessary to formalize "what it means to be a good driver" in a way that is understandable by a human driver. This is an essential point that needs to be addressed in order that society accepts the technology, and the AVs perform well-integrated into a human transportation network.

Q19 How does IEEE P2846 relate to safety assurance and the level of rigor to avoid deviations of defined behavior?

> **Answer by Jack Weast**
> From a technology standpoint, one can formally define what safety means. Models are a good way to do that, and they can be formally verified. Nevertheless, just because it is technically safe, it does not mean that it will be perceived safe by consumers. Drivers or pedestrians may perceive the actions and the behavior of the vehicle to be unsafe and notify the relevant government and industry representatives. This is something the industry has to bear in mind.

4.2 Talk 5: Requirements for Safe Trajectories

Wilfried Steiner
Director TTTech Labs (TTTech Computertechnik AG)

Summary

Self-driving cars generate trajectories to plan their future behavior, but how do we know that these trajectories and the implied vehicle movements are safe? Planned trajectories are internal data structures in the AD system. They mostly consist of a set of waypoints and target velocities. One very first observation is that these data structures need to satisfy some computational sanity checks. For example, whether there are sufficient waypoints present, whether the trajectory's length is within some general bounds. If these sanity checks are satisfied, then the trajectory is possibly safe. However, these sanity checks are certainly insufficient. Some further information we need to reason about the safety of a trajectory is the vehicle's environment, including static and dynamic objects. In general, the planned trajectory is based on how the ego vehicle perceives the world and itself in it. That means the AD system generates a planned trajectory that is based on its world model. However, no model is perfect. Thus, the real trajectory that the vehicle will follow will differ from the planned trajectory. In order to reason about the safety of the planned trajectory, we must also know bounds on the maximum permissible difference of the planned trajectory from the real trajectory the vehicle will maneuver. Still, following these extended trajectory assessments, we cannot be certain if a trajectory is safe. At best, it is possibly safe. This talk argues that one cannot directly determine what constitutes a safe planned trajectory. We can at best determine retrospectively if a real trajectory has been safe. The best we can do is to assess whether a planned trajectory is possibly safe. We need to determine possibly safety indirectly by checks on whether a trajectory is unsafe. Thus, we need to detect a necessary and sufficient set of failure modes. The talk also presents the first set of failure mode classes. Formal models lie RSS already address some of these failure modes. Standardization of these failure modes is essential, and so can be an industry-wide incident response procedure. Such standardization of trajectory safety requirements seems to be complementary to the standardization of requirements on the world model and standardization of requirements on the vehicle dynamics model.

Addressed questions

Q20 How can a safety monitor (i.e., the Checker in Doer/Checker architecture) qualify a trajectory as safe?

> **Answer by Wilfried Steiner**
> The idea behind the safety monitor is to be modular and scalable. Thus allowing different safety verification approaches to be added. Example verification approaches can check for: collision, road departure, excessive planning inaccuracy, vehicle instability, speed limit violation, violation of priority, and others.

Q21 On categorizing behavior as safe/unsafe, should not we consider levels of safety instead of safe/unsafe? How can we do that?

> **Answer by Wilfried Steiner**
> Based on the assumption that the safety monitor will contain multiple verification approaches, it can be that each verification approach provides non-binary results, e.g., a risk value defined as a combination of the probability of occurrence of harm and the severity of that harm. However, in the end, a binary decision, whether the overall AD system's output is safe or not, has to be made. One of the open questions is how the risk assessment of each of the verification approaches can be translated into making a final binary safe or unsafe decision.

Q22 What are sanity checks? How are they performed?

> **Answer by Wilfried Steiner**
> Sanity checks are the most basic form of correctness checks on trajectories. A trajectory essentially is a data structure. This data structure can be verified whether it has a sufficient number of waypoints present, whether the trajectory's length is withing some general bounds, whether it has a too high curvature and others.

Q23 Regarding Safety Co-Pilot: Based on what reasoning is the collision risk threshold chosen and what behavior is carried out when this threshold is exceeded?

> **Answer by Wilfried Steiner**
> The collision risk is an output of a calculation that compares the planned trajectory versus the free-space. The free-space itself is calculated based on the input from the sensor set. Setting the threshold for the collision risk is a complex consideration involving the uncertainties in the previously mentioned calculations and must be further explored. However, the second part of the question is rather more easily answered: when the collision risk threshold is exceeded, then a safety action needs to be performed. In the example of the Doer/Checker architecture with Fallback, the safety action may be to switch over to the Fallback.

4.3 Talk 6: AV Trajectories: Newtonian Mechanics vs. The Real World

Philip Koopman
Prof. at Carnegie Mellon University and CTO of Edge Case Research

Summary

While analysis based upon Newtonian mechanics support proofs of safety for AV trajectory planning, those proofs are subject to assumptions about vehicle capabilities and environmental conditions. Even if those effects are explicitly incorporated into proofs, there will still be measurement uncertainty that needs to be taken into account. Examples of both the types of assumptions and causes of uncertainty are discussed, along with the micro-ODD approach to segmenting and managing proofs of safety within fine-grain ODD subsets.

Addressed questions

Q24 Can the current AD systems determine the road surface state, in order to calculate a proper braking distance for each case? For example, ice, wet road, and others.

Answer by Philip Koopman
The calculations used to estimate the safe following distance of a vehicle are highly dependent on correct environment estimation. Therefore, when developing an AD system, it is important to pay close attention to precise road surface estimation. Simpler systems use a combination of map data that gives baseline road characteristics such as pavement type combined with current weather conditions (rain, snow, dry, and other.). Due to the proprietary nature of AD development, it is unclear if production systems are more sophisticated at, for example, doing real-time estimation of local road surfaces based on sensor data for the road surface in front of the vehicle.

Q25 Vehicles can communicate their intention to break via wireless or so-called vehicle-to-vehicle (V2V) communication. What do you think of this approach?

Answer by Philip Koopman
V2V is good advisory information that can be used to improve the overall AD system safety. However, this approach is not 100% reliable because radio frequency transmission is subject to interference. Therefore, for an example scenario where the following vehicle monitoring for braking by a leading vehicle, complementary approaches should also be used: e.g., observing the front vehicle's brake lights and behavior while ensuring a safe following distance. Ultimately, a safety argument must assume that V2V data might be unavailable. One possible example strategy is always to allow sufficient maximum-deceleration stopping distance in case there are V2V reception faults, but perform more comfortable proactive maneuvers when V2V data is present.

4.4 Panel Discussion on Safe AD Trajectories

Addressed questions

Q26 What is your take for using probabilistic methods for ODD monitoring? "How much" probability is acceptable as a basis for safety decision making?

> **Answer by Jack Weast, Wilfried Steiner and refined by Editor**
> Such methods typically perceive the environment based on a specific set of sensors and reason about the real environment to define in which ODD the vehicle is. The reasoning of the real environment is always a probabilistic best guess, and currently, there are no deterministic methods for guaranteeing correct ODD monitoring - i.e., ODD estimations always have a probabilistic tag to it.
> Therefore, it is crucial to have a proper redundancy in place: i.e., using different types of sensors for estimating where in the world the vehicle is. This may, for example, be a combination of in-vehicle sensors with digital maps and others. Moreover, it is essential to build a proper argumentation that demonstrates the robustness of the ODD estimation method. This argument is then included in the safety case.

Q27 Since monitoring assumptions about other participant's behavior is inherently uncertain: How do you judge the use of probabilistic techniques for RSS monitoring?

> **Answer by Jack Weast and Philip Koopman, and refined by Editor**
> Currently, the best approach for dealing with uncertainty in the prediction of participant's behavior is to work on commonly accepted, industry-wide default set of assumptions. These assumptions will be the baseline to start with, and will likely have to be adjusted for different regional behavioral norms. Later, thanks to innovation and differentiation between technology providers, better probabilistic methods can be developed that will enhance or extend the default set of assumptions. Furthermore, one way to reduce the complexity of modeling the environment is to break up ODDs into pieces (i.e., micro-ODDs). In this way, smaller parts can be proven under assumptions.

Q28 What strategies are useful for transitioning between micro-ODDs?

> **Answer by Editor**
> The speakers were unaware of any strategies for transitioning between micro-ODDs — an indication for relevant research work.

Q29 How can the diagnostic coverage of runtime monitors in AD be measured?

Answer by Jack Weast, Philip Koopman and Wilfried Steiner, and refined by Editor

Classical runtime monitors such as cyclic redundancy checks (CRC), watchdog timers, monitors checking the power supply of the system for under-and-over-voltage, and the like can achieve high diagnostic coverage and are already developed according to ISO 26262. When it comes to monitoring the system's capability or inability to execute a particular function, a new set of challenges is open. For example, "how can one assure implementing a runtime monitor that verifies the AD system's trajectory whether it collides to an obstacle on the road guarantees that a trajectory does not violate other safety properties?". One way to achieve a higher diagnostic coverage of all possible unsafe operations is to look into the safety requirements in the safety case, i.e., the runtime monitor can monitor the validity of these top-level safety requirements, as well as the validity of more detailed assumptions that have been made in the safety case and therefore must be true for the safety case to be valid.

Furthermore, it is expected that the initial runtime monitoring solution will not cover all possible unsafe operations. One way to tackle this issue is to apply a real-time data collection approach that will gather cases where the runtime monitor fails in detecting unsafe operations such as undetected ODD departures or the occurrence of novel edge cases and sends this information up to the cloud. Additionally, runtime collection of Safety Performance Indicators (SPIs) can help inform whether the system is achieving its intended safety targets. Based on this engineering feedback loop, the runtime monitor can be enhanced to cover the missed cases and therefore increase its diagnostic coverage.

Q30 For SAE Level 4 ADS, we expect ODD monitoring from AVs. How robust do you think is the state of technology for ODD monitoring?

Answer by Jack Weast, Philip Koopman, and refined by Editor

Firstly, it is essential to have a common understanding of how an ODD is defined, i.e., is looking into location, weather, and road condition enough, or more is needed? (We think more is probably needed, with UL 4600 section 8.2 [9] enumerating a number of ODD dimensions potentially relevant to safe operation.) An effective way to tackle this problem is by standardizing the ODD definition. Once the ODD is defined, approaches for detecting their characteristics need to be developed. For example, already in place systems like ABS can provide some information about the road condition. Redundant real-time micro weather stations can provide information about the weather. However, common cause failures for the last should not be underestimated, i.e., two micro-weather service providers may receive information from the same thermometer with the same wireless data service that has a shared cell tower, shared backhaul line, shared cloud service, software stack with common vulnerabilities, and others.

5 | Survey Results

5.1 Contributors

In total, 29 contributions were made to the post-event survey. A summary of the contributors' workplace, their role, company/institution, and experience is summarized in Figure 5, under Survey Question (SQ1-SQ4). Contributors' workplace was from 11 different countries. Concerning their current role in the company, the distribution is as follows: 32% general technical/engineering (e.g., system architect, project engineer), 29% general safety (e.g., functional safety, safety experts), 21% research and teaching-oriented (e.g., Ph.D. student, Professor), and 18% have managing roles. Furthermore, 28% are working in a research institution or university, 28% in a Tier 1 company, 21% in a Tier 2 company, 10% for OEM, 3% for a semiconductor company and 3% for other. Finally, 55% of the contributors have less than 5 years, 38% have between 5-10 years, 4% have 10-15 years, and 3% have more than 15 years of experience in the AD domain.

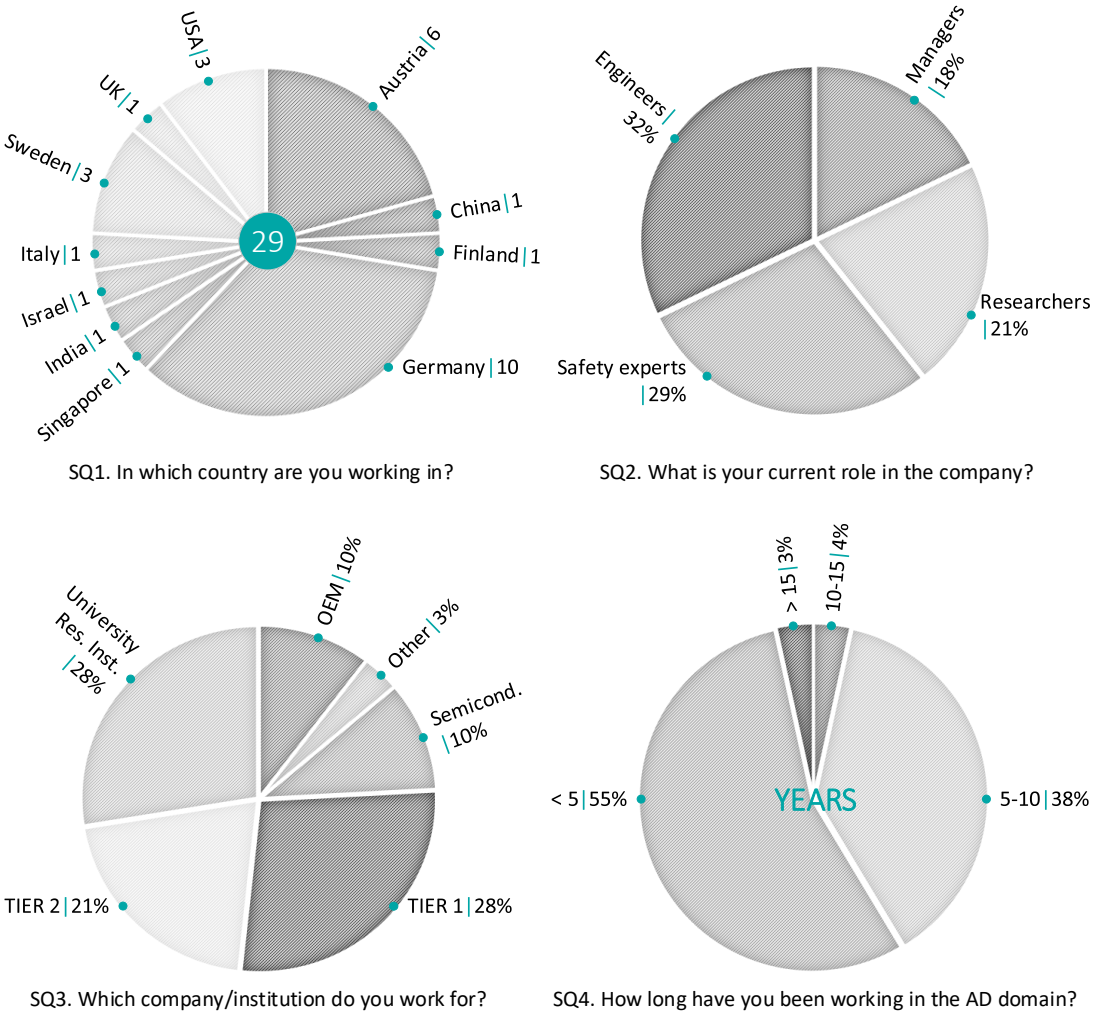


Figure 5: Information about the contributors of the survey.

5.2 Subject: General AD

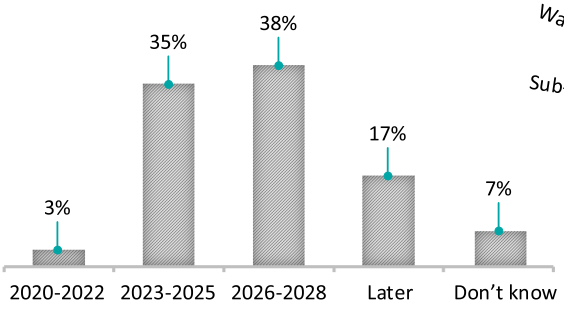
The topics discussed during the Chapter Event often focused on challenges that are faced during the development of future SAE L4 ADS[2]. Therefore it is important to have a common understanding of when SAE L4 ADS are expected to be on public roads and which ODD will come in first. The results of the questions are summarized in Figure 6.

SQ5 When do you expect SAE L4 AD series production vehicles to be on public roads?[3]

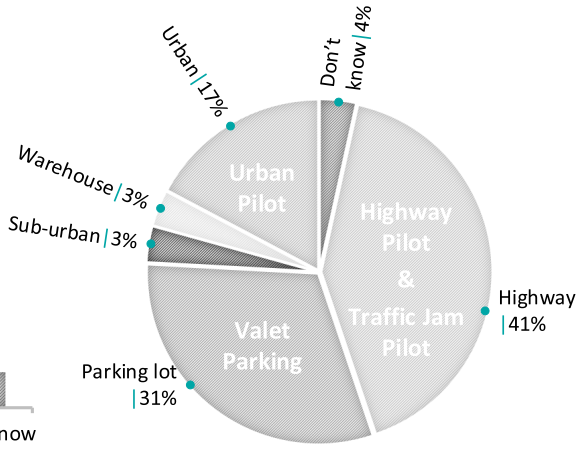
Results
The majority expect SAE L4 AD to be on public roads between 2026-2028 (38%) or between 2023-2025 (35%). Some 17% expect them to be later than 2028, whereas 3% expect them between 2020-2023. Last, 7% did not know.

SQ6 Which ODD do you expect SAE L4 AD series production vehicles to be used in first?

Results
Most of the contributors (41%), expect SAE L4 AD series production vehicles to be first used in Highway related ODD, e.g., highway pilot, traffic jam pilot. Another major part (31%) expect these vehicles on the parking lots, e.g., valet parking. The urban pilot comes third with 17%, whereas sub-urban (e.g., Waymo in Phoenix suburbs [10]) and warehouse (e.g., logistic vehicles) with 3%.



SQ5. When do you expect SAE L4 AD series production vehicle to be on public roads?



SQ6. Which ODD do you expect SAE L4 AD series production vehicles to first be used in?

Figure 6: General AD questions.

[2]SAE L4 ADS performs the complete dynamic driving tasks (DDT) and DDT Fallback (i.e., no fallback-ready driver needed) within a limited ODD.

[3]The editor realizes that the question might have been misleading, as the ODD is not specified. A lesson learned for the next survey.

5.3 Subject: The Autonomous

It important for an initiative to continuously receive feedback from contributors on the selected approaches and vision. Hence, we asked the following question:

SQ7 Do you think the approach proposed by The Autonomous is feasible?

Results

Figure 7 depicts the results. The majority (86%) of the survey participants do believe that The Autonomous approach is feasible, whereas 4% don't. 10% have provided no answer.

For the sake of transparency, opinions (positive and negative) from the survey contributors are summarized below.[4]

Opinions from participants

1. While this approach is feasible, it is not easy and depends to a large degree on the level of detail. Results like the eGas concept document demonstrate that an approach with a suitable, but not to a high level of detail will be accepted and used.

2. A holistic approach is paramount given the complexity of the challenge. Due to legislation in different countries, reference solutions can not be treated globally.

3. For an ecosystem to develop, it needs open and standardized reference implementations. ROS has successfully done that on the framework and middleware level. The Autoware Foundation is currently doing this on the application/and algorithmic level. Unification of the activities can be beneficial.

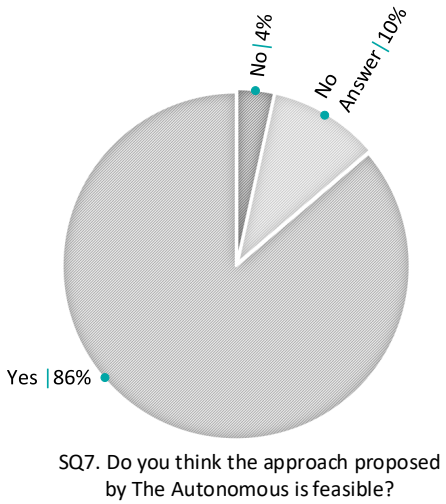


Figure 7: Results from survey question SQ7.

[4]Only spelling and grammar changes have been made. The out-of-context text has been removed.

Opinions from participants

4. It is key to join the forces of the relevant players in this field Reference solutions are a powerful method to prove the state of the art development of AV. But the drawback is the limited competition in terms of technical solutions (but instead on price)

5. Reference architectures that conform to standards is the right approach. Reference scenarios and test cases also need to be developed to demonstrate safety. The main challenges for AVs are the unavailability of the lessons learned, lack of shared data, including best practices, additional to the level of maturity of the safety standards related to AVs. The approach will focus on the gaps for the AV life-cycle in order to define a unified safety framework to cover all the connected AV related aspects. In my opinion, the automobility is waiting for such an approach like The Autonomous.

6. The complementary actions to standardization organization and others are needed and welcome but need to be completed by identifying interdisciplinary topics. The history has shown that the high technical robustness of safety-critical systems is achieved when we have one common authority in defining the rules. The future of human and technical trusted autonomous mobility can reside in The Autonomous initiative.

7. There are so many standards and references to follow for AD, and it makes sense that some organizations gather all and create a full set of requirements to be a unique reference to followed by AD companies.

8. It is feasible to a certain extent. Collaboration is the only way. The Autonomous has the advantage of being more flexible than standardization bodies.

9. The project "Fully Automated Vehicle" is too big to be handled by a single OEM alone. Only through standardization and collaboration can all OEMs and TIERs tackle the technical, legal, and other issues.

SQ8 In your opinion, what do you think the main challenges will be for forming The Autonomous ecosystem?

Results

Figure 8 summarizes the main challenges indicated by the participants. These are:

1. **Openness:** partners may not want to commit and share their know-how before breaking even.
2. **IP rights:** agreement on IP rights is a complex process.
3. **Business model:** creating a business model that fits to everyone's interest is difficult.
4. **Coordinating the partners:** a strong commitment is needed to coordinate the partners in efficient fashion.
5. **Others:** getting disruptors, showing measurable progress, manpower, harmonizing AI and safety, harmonization with standards.

Opinions from participants

1. There are currently many competing standards and organizations trying to accomplish similar things. Only very few of these will remain mid-term.
2. Getting the right key players to participate and showing measurable progress will be a key challenge in building such an ecosystem.
3. Incorporating legislation requirements and consumer needs are different among countries/markets.

Figure 8: Word cloud: challenges for forming The Autonomous ecosystem.

5.4 Subject: Safe AD Architectures

SQ9 In your opinion, what is the key challenge in the development of FO/FD AD architectures?

Results

Figure 9 summarizes the main challenges indicated by the participants. These are:

1. **Costs:** The primary challenge for AD is achieving a reasonable cost while providing extremely high performance. In high price cars, this may not be the key challenge, but certainly a challenge in the mid/low price segment.
2. **Proving safety:** Proving from a regulation point of view that everything reasonably possible has been done to ensure the system's safety will be very hard. This branches out into all areas such as perception and SOTIF.
3. **Real-time capabilities:** real-time data fusion, mapping, decision making, and actuating in complex ODDs.
4. **Risk awareness:** implementing proper "risk awareness" so that the vehicle "understands" the current situation and its current capabilities.
5. **Others:** computing power vs. power consumption, sensor and sensor fusion design maturity, SOTIF, discrimination of anomalies, liability.

Opinion from a participant

The challenge is not the development. It is getting stakeholders to agree on one standard and making that standard open (see ROS which is open, vs. AUTOSAR, which is closed).

Figure 9: Word cloud: challenges in the development of FO/FD architectures.

SQ10 Do you think AD systems architectures should be scalable? i.e., from SAE L2 AD to SAE L5 AD?

Results

Figure 10, left pie chart summarizes the results. A majority (55%) of the survey participants do think that AD systems architectures should not be scalable. Another large portion of the participants (41%) think that the AD systems architectures should be scalable. 4% have provided no answer.

SQ11 Which classic fault-tolerant architecture do you consider most suitable for future FO/FD AD systems?

Results

Here, a significant portion (55%) of the contributors consider that Doer/Checker with Fallback as the most suitable architecture for FO/FD AD systems. Whereas 17% believe that TMR is more suitable. 14% suggest other FO/FD architectures such as 1oo2D or 2oo2D. 14% have provided no answer.

Opinion from participants

More than the selection of fault-tolerant architecture, the proper use case analysis of invoking the appropriate fault-tolerant steps needs to be decided.

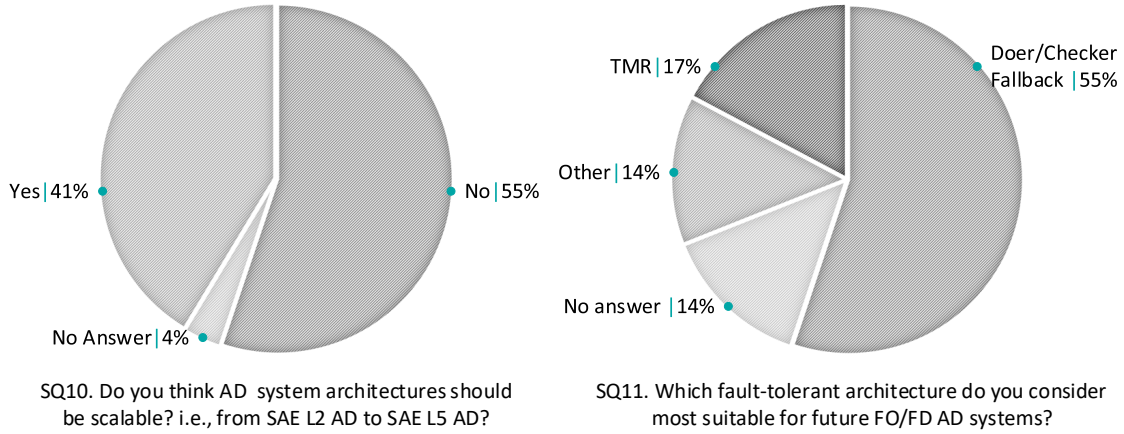


Figure 10: Results from survey question SQ10 and SQ11.

SQ12 Do you think completely independent sensor hardware for each of the redundant channels is necessary to ensure end-to-end redundancy?

Results

Figure 11, right pie depicts the results. The participants, to a large extend (59%), think that there is no need for completely independent sensor hardware for each of the redundant channels. Whereas 38% think that there is a need. 3% have provided no answer.

Participants - justifying their answers

No: I deem sensors much more reliable than L4/L5 decision making. Sensor reliability can likely be achieved without such redundancy.

No: Fault detection and graceful degradation might be enough in some ODDs for some sensors.

No: Hardware-level redundancy is the traditional old school approach. Data integrity is relevant.

No: Dependent fault in sensor hardware can be detected by plausibility checks. Only in cases where this is not possible, the question has to be answered with "yes".

No: From a cost perspective, this seems unlikely. It will be very challenging, though, to come up with a reasonable concept for ensuring sufficient independence between the different channels.

No: Sensor Fusion will provide an independent argument.

No: It can only ensure HW redundancy.

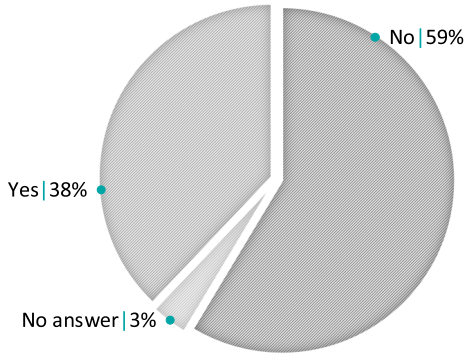


Figure 11: Results from survey question SQ12.

Participants - justifying their answers

No: This depends on the system fault impact and response due to the loss of a sensor. Furthermore it depends on the ability to detect the faulty sensor.

No: Some of the sensors could be common, but then some complementary fail-operational sensors could be used. The fail-operational control mechanisms should be able to operate with a limited sensor input, and the complementary sensor HW should be able to support even the fail-operational state.

No: A part of common sensors can be used in combination with independent sensor technology.

No: If you mean two separate channels completely redundant and diverse, I disagree since if any fault happens in one of these two separate channels then we will lose the whole chain and it is not efficient (e.g., by loosing just a processor we might lose the whole chain and diversity of sensors).

No: Depends on the definition of end-to-end redundancy, but a common cause failure is acceptable if it can detect and handle it or if it is extremely unlikely and has no severe consequences.

No: Sensor aggregation over time, costs.

Yes: A non-independent sensor would not allow end-to-end redundancy by definition. If that sensor fails, both channels fail.

Yes: It would be better to avoid common systematic failures.

Yes: To eliminate common cause failures.

Yes: To prevent common-mode failures and common cause failures, it will be more suitable to use different technologies.

Yes: If we have heterogeneous systems (including multiple sensor hardware), we have a lower probability of failure caused by the exact same fault.

Yes: Taking into account the maturity of the systems, we should rely on completely independent hardware (at least for now for R&D trials). As the maturity of the AD system increases and the functionality of the AD system is proven to be stable (for XX years), we can optimize the system in terms of functionality and efficiency.

SQ13 What best practice would you recommend in the development of the FO/FD AD system architectures.

Results

Due to the novelty of the field under study, not many practices exist at the moment. Figure 12 summarizes the recommendations given by the participants.

1. **Best practices from other domains:** Studying best practices from other domains is recommended as a good starting point. Aerospace/avionics (e.g., ARP4761 [11]) is one relevant domain to consider looking into. Certainly, one needs to have (i) a good understanding of which circumstances make the architecture a best practice in the other domain and (ii) tailor the solution to the automotive use-case (if possible).
2. **Lessons learned:** *"Because those who have not read history are doomed to repeat it"*. Bad experiences from other domains should also be studied and considered to avoid repeating mistakes. A comprehensive list can be found in [12].
3. **Algorithmic redundancy:** take advantage of algorithmic redundancy and make explicit under which calculation assumptions an algorithm provides correct information. Consider these assumptions as a driver for the selection of algorithms at runtime.
4. **Modularity:** Designing a modular system will enable easy reconfiguration of the system.
5. **Testing at night:**"Nightly" automated testing with very high coverage on the integration level, i.e., car-in-the-loop in virtual reality based on real data (captured during real driving) and using (emulated) failure injection.

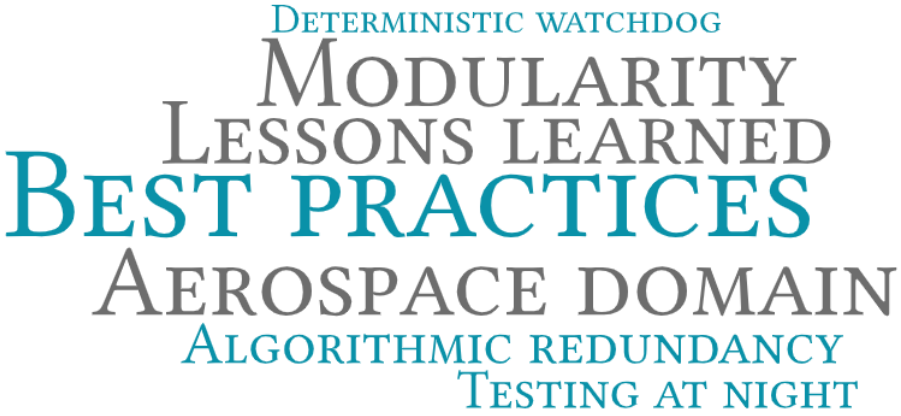


Figure 12: Word cloud: best practices and recommendations for the development of FO/FD AD system architectures.

SQ14 An AD system is expected to receive continuous updates after being deployed on public roads in order to improve functionality. How can shorter update cycles be guaranteed without jeopardizing the overall system's safety?

Results

Figure 13 summarizes the recommendations given by the participants.

1. **V&V:** appropriate verification and validation before being pushed to the fleet.
2. **Simulation:** Realistic simulation x-in-the-loop-based regression testing before release and dual opinion between releases in the vehicle side.
3. **Modularization:** Separating the functional from the safety-critical part can reduce the risk of potential safety issues. In this way, the functional part can be updated as often a needed.
4. **Digital twins:** deploy platforms that predict the behavior of these updates at runtime. For example, through digital twins - a virtual evaluation that runs faster than the wall clock.
5. **Traceability and evaluation metrics:** The AD system needs traceability, which allows a proper impact analysis whenever there is an update. This, coupled with critical evaluation metrics for the system, could provide confidence.
6. **Clear item and requirements definition:** clear item and requirement definition are necessary to avoid additional software updates because of misunderstood or missed requirements/problems.
7. **Others:** formulized modular safety cases, sandboxing concepts.

Figure 13: Word cloud: best practices and recommendations for achieving shot-update cycles without jeopardizing the overall system's safety.

Opinion from a participant: The security perspective
By establishing a robust security approach that evolves along with the potential threats. A vulnerability analysis and associated mitigation approaches need to be considered upfront during the architecture development, along with approaches to update them as new threats occur. Approaches such as encryption, software signature checks, virus/malware detection, anti-virus updates, and others. can be used. A safety interlock approach that only allows an update to occur after certain predefined conditions could be used (similar to aerospace). Also, a statically configured, table-driven approach for critical functions could be used, including checking versions or other compatibility checks to vote out non-conforming elements that may not be adequately updated or have been corrupted. With V2X, viruses could be passed from vehicle to vehicle. This area should get much attention to threat assessment and mitigation approaches.

5.5 Subject: Safe AD Trajectories

SQ15 What are the criteria that an AD system's output (i.e., a trajectory, path) should meet to ensure that it is safe?

Results

Figure 14 depicts the results. Most of the participants (36%) think that an AD system's output should comply with all four criteria, i.e., (i) not colliding with static or dynamic obstacles, (ii) not leaving the road edge, (iii) not leading to loss of stability, and (iv) not breaking the legal rules. 24% have voted for not colliding to obstacles on the road, 17% for not leading to loss of stability, 9% to not leaving the road edge, and 2% to not braking the legal road rules. 7% have proposed other criteria, such as avoid the AD system's output outside the ODD, staying in lane, and assuring the safety of vulnerable road users (especially for Urban ODD). 5% of the participants have provided no answer.

Opinions from participants

1. The criteria should be prioritized and sometimes not followed for the sake of higher safety goals: e.g., in order to be able to avoid an obstacle, the not leaving the road edge rule has to be ignored.
2. The criteria are neither necessary nor sufficient conditions and might be conflicting in some situations, but it is a reasonable minimum set to keep in mind for engineering.
3. One should leave sufficient safety margins to account for uncertain/incomplete information/disturbances.
4. Road rules can be contradictory, and until they are written for robots, they can not have higher priority than other goals.

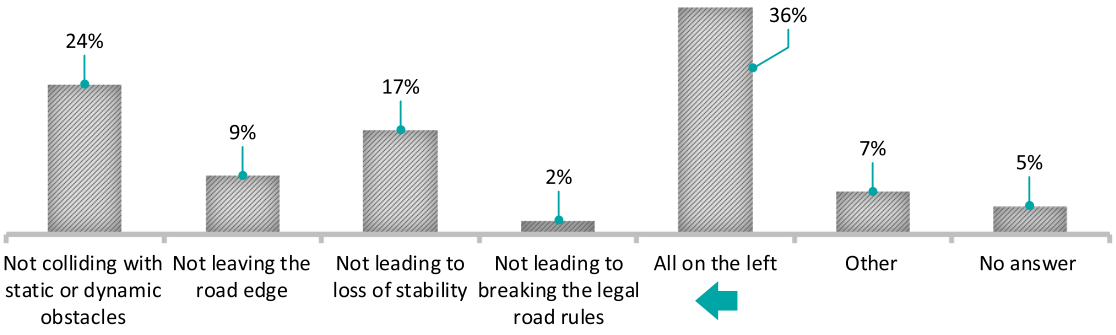


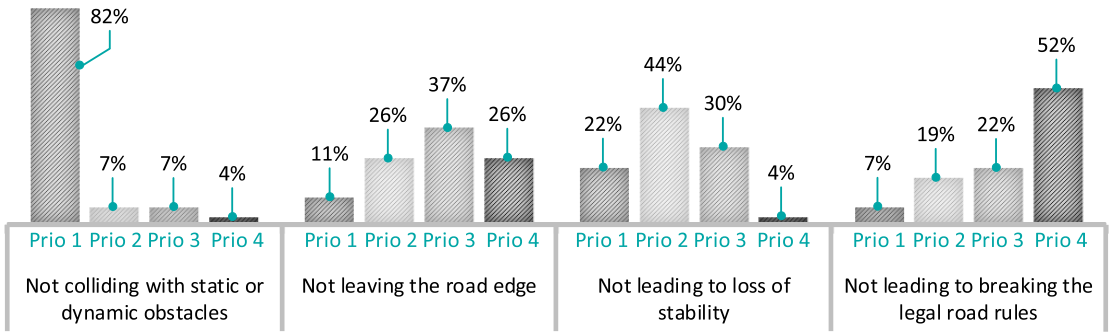
Figure 14: Results from survey question SQ15.

SQ16 How would you rank the criteria defining a safe AD system's output (i.e., a trajectory, path), e.g., Prio 1 (highest priority) Prio 4 (lowest priority)?

Results

Figure 15 depicts the results.

- **Not colliding to static or dynamic obstacles:**
  - 82% Prio 1,
  - 7% Prio 2,
  - 7% Prio 3,
  - 4% Prio 4.
- **Not leaving the road edge:**
  - 37% Prio 3,
  - 26% Prio 2,
  - 26% Prio 4,
  - 11% Prio 1.
- **Not leading to loss of stability:**
  - 44% Prio 2,
  - 30% Prio 3,
  - 22% Prio 1,
  - 4% Prio 4.
- **Not breaking the legal road rules:**
  - 52% Prio 4,
  - 22% Prio 3,
  - 19% Prio 2,
  - 7% Prio 1.



SQ16. Rank the criteria defining safe AD system's output (i.e., a trajectory, path) by their importance.

Figure 15: Results from survey question SQ16.

SQ17 Do you think common/standardized interface definitions for AD system's output (e.g., trajectory, path) should be defined?

Results

Figure 16 depicts the results. A major portion (73%) think that there is a need for a common/standardized interface for AD trajectories. Whereas, 17% have answered with "No", and 10% have provided no answer.

Participants - justifying their answers

No: I don't believe in setting a standard before we know what practice is the best.

No: First, the industry needs solutions. Standards before solutions could stop innovations

Yes: Standardized interfaces are a potential approach to ensure interoperability.

Yes: To ensure reproducible and calculable behavior, and to enforce a joint effort to understand the challenge and find a reliable solution.

Yes: Allows contribution from multiple suppliers/vendors to the same system.

Yes: Standardization enables interoperability.

Yes: This could facilitate modularization of the system and also make establishing a state-of-the-art easier as no apples-to-oranges comparisons would be necessary.

Yes: It facilitates standardization and reduces the cost for collaboration with suppliers.

Yes: This would enable the inter-interoperability of technologies developed in this sense. Furthermore, it would enable enhancements in distinct directions while assuring the interoperability.

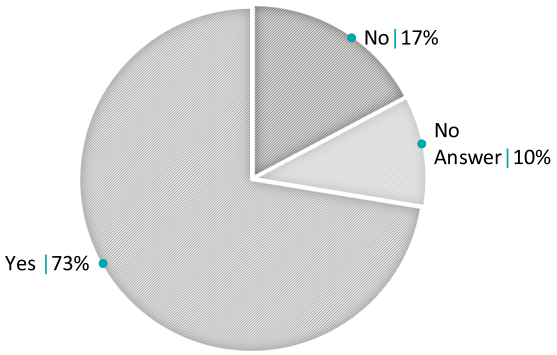


Figure 16: Results from survey question SQ17.

Participants - justifying their answers

Yes: I am not sure if standardization is desired, but common interface definitions would be a great benefit. At least some guidelines for an interface that a system could be evaluated against are needed. Standardized may be too limiting, but a deeper discussion is needed.

Yes: A standardized interface shall increase the chance for the different stakeholders in AD space to work together towards a common goal: i.e., provide AD system's safety - AUTOSAR is a good example.

Yes: Starting from the common ground makes it possible to converge. If manufacturers use different definitions, I don't know how there can be interoperability and good explainability.

Yes: The task of building the architecture for highly AD is huge, so a certain level of standardization of common interfaces/tasks/methods could save time, money, and make life easier for all of us.

Appendices

A | List of Abbreviations

AD	Automated Driving
ADAS	Advanced Driving Assistance Systems
ADS	Automated Driving System
AI	Artificial Intelligence
ANSI	American National Standards Institute
ASIL	Automotive Safety Integrity Level
AUTOSAR	Automotive Open System Architecture
AV	Automated Vehicle
CD	Commission Draft
CPS	Cyber-Physical System
ECU	Electronic Control Unit
FO/FD	Fail-Operational/Fail-Degraded
FuSa	Functional Safety
ISO	International Standardization Organization
L1	SAE Level 1
L2	SAE Level 2
L3	SAE Level 3
L4	SAE Level 4
L5	SAE Level 5
ODD	Operational Design Domain
OEM	Original Equipment Manufacturer
PAS	Publicly Available Specification
SAE	Society of Automotive Engineers
SaFAD	Safety First for Automated Driving
SOTIF	Safety of The Intended Functionality
TR	Technical Report
UL	Underwriters Laboratories
V&V	Verification and Validation

B | Compliance Guidelines

Ensuring safety is the key to gaining acceptance of autonomous mobility on a broad scale. The Autonomous will start this critical discussion by gathering together the complete autonomous mobility ecosystem and facilitate a mutual exchange of ideas by offering various workshops on key topics (Safety & Security, Safety & AI, Safety & Architecture, Safety & Regulation), panel discussions, and keynote speeches.
At The Autonomous, we are committed to ensuring that all discussions take place in full compliance with the rules of competition law. In order to allow for an open exchange of ideas within the limits of the law, this Guideline sets out practicable rules for The Autonomous. Compliance with this Guideline is obligatory for all organizers and participants.

1. **Permitted topics:** Topics which may be covered in discussions, workshops and meetings organized by The Autonomous include:

    1.1. General technical and scientific developments relevant to mous mobility;

    1.2. Legislative proposals and/or regulatory measures and their impact on the autonomous mobility ecosystem;

    1.3. The political environment;

    1.4. Current economic developments and general developments in the industry (if publicly available);

    1.5. Exchange of freely available information e.g. economic data available online or in annual reports.

2. **Non-permitted topics:** Participants may not discuss, agree, share information on, or in any other way coordinate their behavior regarding competitively sensitive issues, including:

    2.1. Current and future prices, including selling prices, purchase prices, price components, price calculation, rebates, and intended changes in prices;

    2.2. Terms and conditions of supply and payment for contracts with third parties;

    2.3. Market sharing, including discussions on the division of sales territories or customers (e.g., by size, product type, etc.);

    2.4. Co-ordination of bidding towards third parties, including information on customers' commercial expectations and the firm's proposed response, as well as information on proposed bids (whether a bid will be submitted, for which lots, etc.);

    2.5. Boycotts against certain companies, e.g., agreements not to work with certain customers or suppliers, or to exclude specific companies from discussions on the establishment of a technical standard;

    2.6. Information about business strategies and future market conduct, such as planned investments or the commercial launch of new technologies or products (if not publicly available). In particular, agreements to delay a new technology or to fix the commercial terms of its introduction are prohibited;

2.7. Detailed information on financial performance, such as recent information on profits and profit margins on a non-aggregated basis (if not publicly available);

2.8. Information on internal research and development projects. This comprises estimations about the feasibility of specific technical solutions or the costs attached to the implementation of a specific solution.

3. **Measure to ensure compliance:** In order to ensure compliance and to contribute to an open discussion, The Autonomous will implement the following measures:

   3.1. Attendance by legal counsel: All discussions and workshops will be attended by in-house or external legal counsel. Legal counsel may break off or adjourn the discussion in case of doubts with regard to competition law compliance.

4. **No Reliance:** The purpose of this Guideline is to briefly summarize the competition rules applying to discussions at The Autonomous. It, however, cannot address the full complexity of the applicable law and does not constitute legal advice to participants and their respective firms as to their obligations under competition law. At The Autonomous, we encourage participants to familiarize themselves with the rules of competition law. Should any participant have doubts as to the legality of any discussion in the course of The Autonomous, she/he may:

   4.1. raise such doubts to the legal counsel attending the discussion. The legal counsel shall record any such request in the minutes;

   4.2. leave the meeting if the discussion continues without the participant's doubts having been resolved. The legal counsel shall record the name of the participant as well as the exact time of the participant's departure in the minutes.

C | Standard Settings Guideline

Ensuring safety is the key to gaining acceptance of autonomous mobility on a broad scale. To address security concerns in connection with autonomous driving, safety proves to be the main concern and challenge for mass adoption. These current challenges and associated investment costs cannot be mastered by a single OEM, Tier 1, or Tech company. Just like in aviation, autonomous driving needs to set common technical and ethical standards, legislation, and a process to learn from past incidents and avoid future ones.
At The Autonomous, our mission is to establish a global safety reference, created by the global community, which facilitates the adoption of autonomous mobility on a grand scale. We are committed to ensuring that this process takes place in full compliance with the rules of competition law. To this end, this Guideline supplements The Autonomous' Compliance Guideline, by setting out practicable rules for standard-setting processes at The Autonomous. Compliance with this Guideline is obligatory for all organizers and participants.

1. **Openness and transparency:** The Autonomous follows an open and transparent approach to participation in its panels, workshops, and other working groups. The establishment of a global safety reference will follow the following principles:

    1.1. Unrestricted participation: involvement is open to all industry stakeholders. Active involvement may only be limited if absolutely necessary (i.e., to prevent inefficiency) and based on objective and non-discriminatory criteria;

    1.2. Transparency: all attendees of The Autonomous, as well as all other stakeholders concerned, will be informed of any announcement, progress, and outcome;

    1.3. Review and comments: Stakeholders not participating in the process will be able to review and comment on the result of the standard-setting process. Any agenda referring to activities of The Autonomous will be disseminated to participants in due course prior to the execution of the activity. Participants shall have the right to comment or to contribute to such an agenda.

2. **Non-exclusivity, free access**

    2.1. No obligation to comply: Participants are free to develop alternative standards or products that do not comply with the evolving standard;

    2.2. Free access to standards: Any developed standards will be accessible for all interested stakeholders (whether or not they participated in The Autonomous) on fair, reasonable, and non-discriminatory terms.

3. **IPR Policy**

    3.1. **Definitions:**

        3.1.1. "Affiliate": any subsidiary or holding company of a participant, any subsidiary of any of its holding companies and any partnership, company, or undertaking (whether incorporated or unincorporated) in which a participant has the majority of the voting rights or economic interest.

3.1.2. "Essential": an intellectual property right is essential where it would be technically (but not necessarily commercially)impossible, taking into account normal technical practice and state of the art generally available at the time of adoption of the standard, to implement the respective standard without making use or infringing the IPR in question.

3.1.3. "FRAND terms": fair, reasonable, and non-discriminatory terms.

3.1.4. "Implement/Implementation": (i) to make, market, sell, license, lease, otherwise dispose or make use of equipment; (ii) repair, use or operate equipment; or (iii) use methods – as specified in the respective standard.

3.1.5. "Intellectual Property Rights" or "IPR": any copyright, Patent, registered design, and any application thereof. IPR does not include trademarks, trade secrets, moral rights, right of know-how, and confidential information.

3.1.6. "Patent": any patent, utility model, or any application for such.

3.2. **Scope of Application:** Participants owning any Essential IPR shall be free to exploit such IPR outside the scope of The Autonomous at their absolute discretion and any revenues or other benefits, which the participant may receive from such exploitation of such Essential IPR, shall be for the participant's own account.

3.3. **FRAND commitment**

3.3.1. Save in the case of any Essential Patents identified in accordance with Section 3.4.4, a participant will give an undertaking that it is prepared to grant licences to anyone wishing to Implement the standard to which the Essential IPR relates:

(i) on FRAND terms;

(ii) to all its Essential IPR relevant for the respective standard;

(iii) to the extent necessary to permit the Implementation of the respective standard.

3.3.2. The undertaking pursuant to Section3.3.1 may be made subject to the condition that those who seek licenses agree to reciprocate.

3.3.3. Where a participant has elected not to declare or has failed to declare any Essential IPR for a given standard in accordance with Section 3.4.4, the participant shall be deemed to have given the undertaking in accordance with the terms of Section 3.3.1.

3.3.4. Both, the participant who has given an undertaking pursuant to Section 3.3.1 or who is deemed to have given an undertaking pursuant to Section3.3.3, and any beneficiaries of such undertaking wishing to acquire a license in accordance with Section 3.3.1, acknowledge and agree that:

(i) They will act in good faith, in order to negotiate a license agreement;

(ii) If both parties have not been able to agree on an Essential IPR license, each party has the right to pursue the matter before the national courts to resolve the matter.

3.3.5. Each participant will ensure that its Affiliates and its Affiliates' successors in title will give an undertaking pursuant to Sections 3.3.1 to 3.3.4 above. If a participant or its Affiliate transfers ownership of Essential IPR that

is subject to an undertaking 3 pursuant to Sections 3.3.1 to 3.3.4 above, such undertaking shall include appropriate provisions in the relevant transfer documents to ensure that the undertaking is binding on the transferee and that the transferee will similarly include appropriate provisions in the event of future transfers with the goal of binding all successors-in-interest. The undertaking shall be interpreted as binding on successors-in-interest regardless of whether such provisions are included in the relevant transfer documents.

3.4. **Declaration of Essential IPRs**

3.4.1. Prior to any official adoption of any standard or part thereof, each participant shall provide a written declaration of the Essential IPR relevant to the subject matter. Such declaration shall list:

(i) all potentially relevant Essential IPR held by the participant or its Affiliates;

(ii) filing and registration number, application date and if published the title of the respective Essential IPR;

(iii) terms (i.e., explicitly (non-FRAND terms as opposed to clause 3.3.1, but without specifying royalty rates on any other royalty terms)) on which the participant or its Affiliate is prepared to grant licenses to other participants or any third parties; and

(iv) statement whether the declaration is made subject to the condition that those who seek licenses agree to reciprocate.

3.4.2. In the absence of a declaration of any Essential IPR, the participant will be deemed to have given the undertaking for that Essential IPR associated with the relevant standard or part thereof, in accordance with Section 3.3.3.

3.4.3. Any declaration may identify such Essential Patents, for which the participant or its Affiliate are unwilling or unable to enter into an undertaking to license on FRAND terms in accordance with Section 3.3.1. The declaration shall:

(i) identify any such any Essential Patent, by way of filing number, date, and if published, optionally its title;

(ii) describe in sufficient detail the reasons why the participant or its Affiliate are unwilling or unable to enter into an undertaking to license on FRAND terms in accordance with Section 3.3.1.

3.4.4. Where a participant, in accordance with Clause 3.4.3, has identified an Essential Patent, which the participant, or its Affiliates, is unwilling or unable to license in accordance with Clause 3.3.1, the participant will lose its right to participate and to receive undertakings pursuant to Clause 3.3.1 from other participants in relation to the respective standard or part thereof to which an Essential Patent relates, if:

(i) any other participant informs the Chairman within a reasonable period, in writing, that it does not accept that the reasons in the relevant declaration (as required in accordance with Clause 3.4.3(ii))are reasonable and justified; and

(ii) based on its duly justified non-acceptance of these reasons pursuant to Clause 3.4.4.(i), wishes that the aforesaid participant shall not be able to rely on its right to participate and to receive undertakings pursuant to Clause 3.3.1 from other participants.

3.5. **Disputes concerning ownership of Essential IPR:**If two or more participants claims ownership of the same Essential IPR, the participants claiming ownership shall:

(i) negotiate and resolve the question of ownership in good faith and

(ii) if no solution is found pursuant to section s3.5.1, have the right to pursue the matter before the national courts to resolve the dispute.

D | Acknowledgments

First and foremost, sincere thanks to all keynote speakers, namely Jack Weast, Martin Törngren, Philip Koopman, Riccardo Mariani, Simon Fürst, and Wilfried Steiner. Their constant support over the past months and in-depth knowledge in the field resulted in outstanding presentations and discussions.

Furthermore, profound gratitude to all the participants at the virtual Chapter Event as well. Their questions enriched and deepened the discussions throughout the workshop.

Special thanks also go to the contributors of the post-event survey who additionally enhanced the quality of discussions and ultimately of this report. In this post-event survey, the contributors were given the option to select whether their names should be mentioned or not. The following is a list of a substantial number of contributors: Abhash Das, Ali Nouri, Andrea Bondavalli, Chithra C. S., Christoph Schmittner, Christopher Temple, Emilia Cioroaica, Guillermo Rodriguez-Navas, Hendrik Weppelmann, Jan Becker, Jens Rosenbusch, Kenneth Rosol, Kholoud Hatem, Mohamed Azhar, Patricia La Torre, Rasmus Adler, Roman Benesch, Rudi Grave, Sami Dahlman, Sandor Mathe, Truls Nyberg, and Xinhai Zhang.

Likewise, warm thanks to all reviewers - for all your comments and ideas for enhancement you have proposed.

Sincere thanks to Wilfried Steiner, who served as the leading “sparring partner” when writing the report.

Many thanks to Georg Kopetz, Marc Lang, Ricky Hudi, and Stefan Poledna for initiating The Autonomous and believing in this cause.

Last but not least, warmest thanks to The Autonomous team - Iulia Alina Baidac, Luisa Griesmayer, Susanne Blum, and Philip Schreiner - for your excellent work and continuous support.

E | Feedback

In our continuous effort to develop The Autonomous as an open platform and space for dialogue among different stakeholders, we welcome all feedback and interest in making safe autonomous mobility a reality. We highly value any comments, ideas, or suggestions you may have to help improve the outcome of this report or contribute to the initiative. Please do not hesitate to contact us at: [*contact@the-autonomous.com*].

References

[1] International Organization for Standardization (ISO). Road vehicles-functional safety standard, ISO 26262. *International standard*, 2018.

[2] International Organization for Standardization (ISO). ISO/PAS 21448, Road vehicles — safety of the intended functionality. https://www.iso.org/standard/70939.html. Accessed: May-2020.

[3] International Organization for Standardization (ISO). ISO/CD TR 4804, Road vehicles — safety and cybersecurity for automated driving systems — design, verification and validation methods. https://www.iso.org/standard/80363.html. Accessed: May-2020.

[4] BMW. Safety assessment report: SAE Level 3 ADS. https://www.bmwusa.com/content/dam/bmwusa/innovation-campaign/autonomous/BMW-Safety-Assessment-Report.pdf. Accessed: May-2020.

[5] IEEE. P2846: A formal model for safety considerations in automated vehicle decision making. https://sagroups.ieee.org/2846/. Accessed: May-2020.

[6] IEEE. P2851: Exchange/interoperability format for safety analysis and safety verification of ip, soc and mixed signal ics. https://sagroups.ieee.org/2851/. Accessed: May-2020.

[7] Yue Kang, Hang Yin, and Christian Berger. Test your self-driving algorithm: An overview of publicly available driving datasets and virtual testing environments. *IEEE Transactions on Intelligent Vehicles*, 4(2):171–185, 2019.

[8] Junko Yoshida. Autonomous vehicles safety ventures beyond ISO 26262. https://www.eetimes.com/av-safety-ventures-beyond-iso-26262. Accessed: May-2020.

[9] Underwriters Laboratories. ANSI/UL 4600 standard for safety for the evaluation of autonomous products. https://ul.org/UL4600. Accessed: May-2020.

[10] The Verge. Waymo's driverless car: Ghost-riding in the back seat of a robot taxi. https://www.theverge.com/2019/12/9/21000085/waymo-fully-driverless-car-self-driving-ride-hail-service-phoenix-arizona. Accessed: May-2020.

[11] ARP4761, SAE. Guidelines and methods for conducting the safety assessment process on civil airborne systems and equipment. *London-UK: SAE*, 1996.

[12] Philip Koopman. Computer-based system safety essential reading list. https://safeautonomy.blogspot.com/p/safe-autonomy.html. Accessed: May-2020.