SIEMENS
Ingenuity for life

Siemens Digital Industries Software

# Delivering safe automated driving systems

## Identifying trends in testing and certification for autonomous vehicles targeting automation level 4


### Executive summary

In the past decade, significant technological progress has been made in the field of automated driving systems. Despite this technological progress, autonomous vehicles with level (L) 3 automation or higher are not yet on the public roads. The main reasons behind this are the inability to prove safety compliance, meet regulatory requirements and achieve user acceptance. In this white paper we will focus on the first two reasons, highlighting the state-of-the-art and future trends in ensuring safety for automated driving systems.



Alexandru Forrai, Ph.D., and Matthieu Worm,
Siemens Digital Industries Software

# Contents

# Abstract

In the past decade, automated driving systems, advanced driver assistance systems (ADAS) and autonomous vehicles (AV) were the focus of intensive research and development (R&D). Despite the significant progress made in the case of ADAS, for AVs the following three main challenges still exist:

Technology challenge: Build a safe car, which means it can perceive the road environment better than a human driver and makes reasonable decisions like a human driver.

Regulatory challenge: Build a functional car accepted by society, which means it makes a proper tradeoff between safety and functionality: "I am safe if I do not drive, but then I am not useful." It fits into the defined regulatory framework (for example, testing and certification).

Business challenge: Build a cost-effective car, which means customers are willing to switch to driverless cars and/or to new business models (for example, redefinition of mobility).

Since automated driving systems are highly complex systems in terms of hardware and software, these challenges are not independent so there is a strong interaction between them.

For example, a technical and regulatory challenge can be: how to prove safety compliance and meet regulatory requirements? A possible or obvious answer is: high complexity requires massive verification and validation (V&V) cycles. Thus, it is foreseen the amount of physical testing and virtual testing will increase, which will impact the costs. Therefore, a technical and regulatory challenge is strongly linked with a business challenge.

On the other hand, these challenges are not AV-specific, they are general challenges, which different industry sectors face. For example, how to build a safe airplane, make air traffic safe or build a safe oil refinery. The good news is other industry sectors can address these challenges properly and can be a good inspiration for the automotive industry as well. For sure, the automotive industry will progress to a higher level of driving automation in the coming years and developed technologies will also impact other industry sectors.

# Ensuring safety for automated driving systems: relevant standards

Let us start with the regulatory challenges: What is safety and how can we define it? Safety has many definitions, for example: freedom from conditions that can cause death, injury, occupational illness, damage to or loss of equipment or property, or damage to the environment, according to the military standard MIL-STD-882E.

Next, if we narrow down the application field to passenger vehicles and focus on operational safety, we can observe that safety has many aspects, as shown in figure 1.

Safety aspects are covered by different standards as shown in figure 2; some of them are region-specific (for example, traffic rules) and some of them are more general and are widely used by the automotive industry (for example, International Organization for Standardization 26262). From an AV perspective, we can state that despite significant work and effort currently there is no single, unified standard that might provide guidance during development and certification.[1]

Despite this, two relevant standards – ISO26262 and safety of the intended functionality (SOTIF), which complement each other – play a key role in the development process of automated driving systems.[2,3]

ISO26262 is the functional safety standard that specifies how the system should detect and respond to failures, errors, or off-nominal performance.[2]

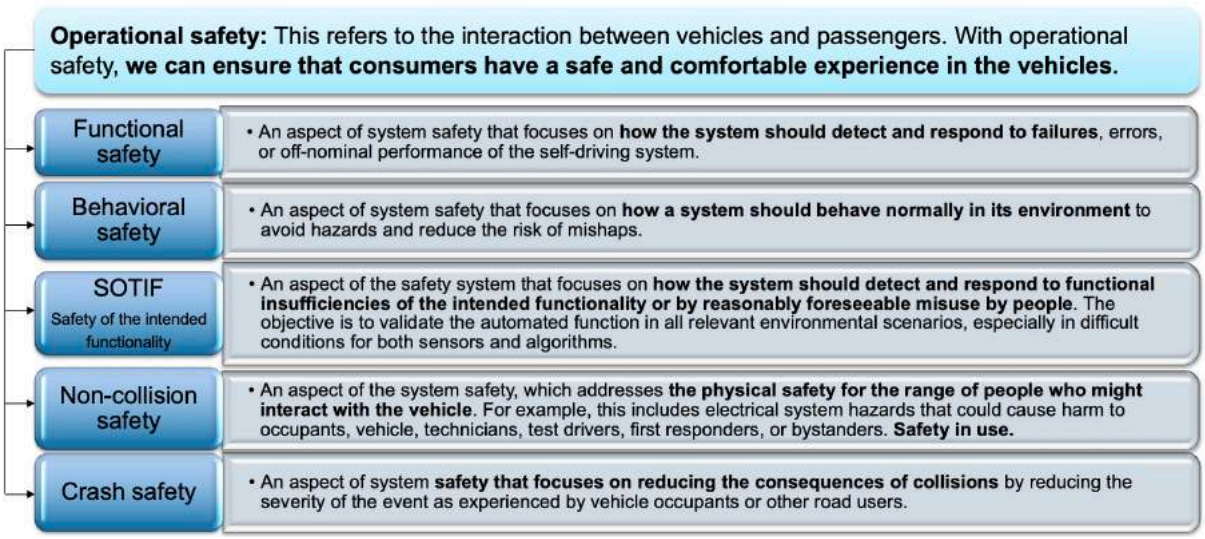


Figure 1. Different aspects of operational safety.

SOTIF specifies how the system should detect and respond to functional insufficiencies of the intended functionality or by reasonably foreseeable misuse by people.[3]

From a safety point of view, the main objective of the development is to properly assess and take proper measures to minimize risk at least to the accepted (tolerable) level of risk, as shown in figure 3, where the accepted risk is defined by the society, which can be region-specific.[4]

So how we are going to reduce the risk? From a functional safety perspective (ISO26262), risk is due to hazards caused by malfunctioning behavior of electric/electronic and programmable systems (see figure 4).

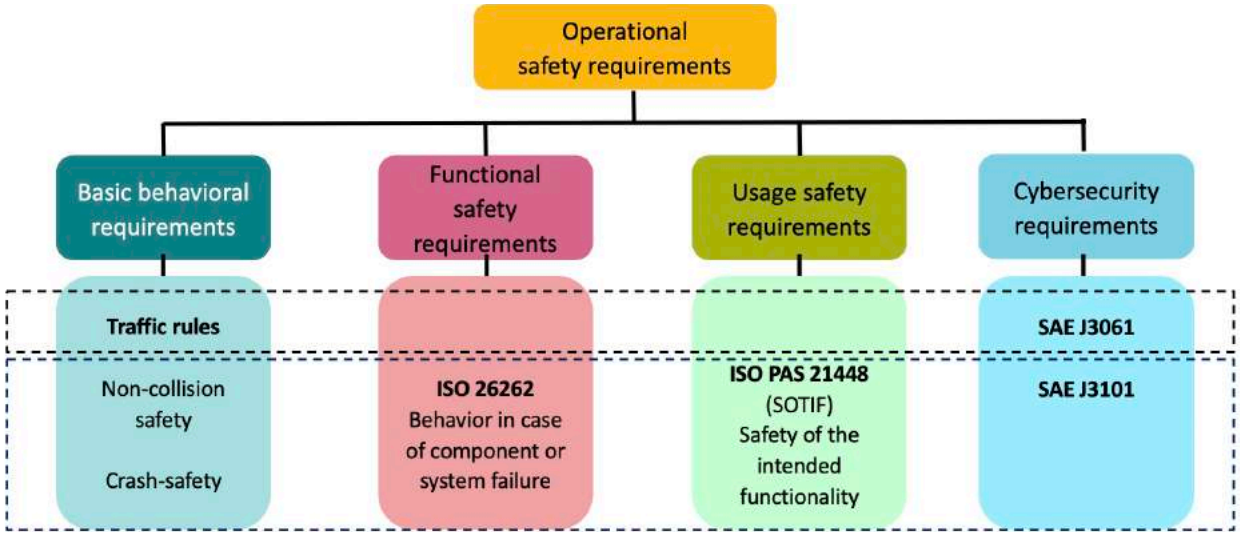


Figure 2. Relevant safety standards for automated driving systems.

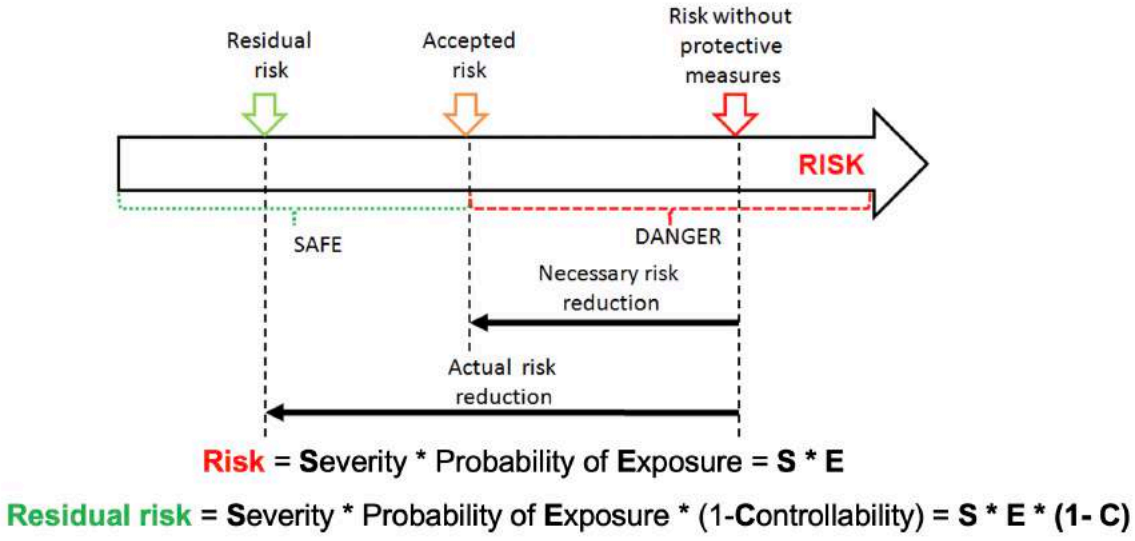


Figure 3. Risk and residual risk.

The failures within electric, electronic and programmable systems are divided into two categories:

- Systematic failures – such as software bugs, faults introduced during hardware design and faults introduced or not detected by development tools
- Random hardware failures due to component failures because of material imperfections, manufacturing, aging, etc.

Reducing the risks due to hardware failures is done by increasing the area of dangerous detected faults and decreasing the area of dangerous undetected faults (see figure 5). In practice, the area of dangerous undetected faults is minimized with diagnostics, redundancy and better-quality components.[4]

How much risk reduction is required or what is the required safety and integrity level depends on how safety-related or safety-critical is that system or subsystem. The Automotive Safety Integrity Level (ASIL) sets the standard for the industry. For example, the required safety level for an electric power steering system or the deployment of the airbag system is ASIL D (highest automotive safety integrity level).

Reducing the risks due to systematic failures, such as software bugs, faults introduced during hardware design and tools, is done by following a well-defined development process. One important approach to reducing risk due to systematic failures is virtual verification and validation, which we will focus on later.


Figure 4. Functional safety standards used in different industry sectors.

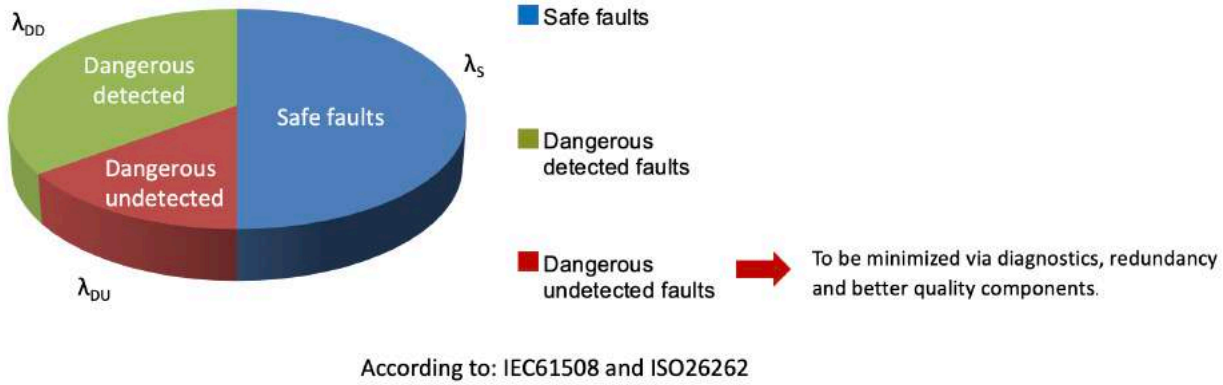


Figure 5. Reducing the risks due to hardware failure.

At this stage, let us consider a simple example: The object detection and classification algorithm of an automated driving system is going to be verified and validated using different images, with and without faults, as shown in figure 6. It might be the object detection and classification algorithm is not able to classify the traffic sign correctly when faults are injected; in such a case, redundancy/ diversity is a possible technical solution: for example, usage of high definition (HD) maps, which have traffic signs embedded.

Next, we might ask how about the faults introduced during development by the tools which we use? The latest edition of the ISO26262:2018 standard recommends qualifying tools used in developing safety-related systems. A tool might introduce errors during the design process or might be unable to detect errors during the verification and validation phase. Therefore, in the first stage the tool confidence level is assessed (based on tool impact and error detection level) and if the confidence level is low, a so-called tool validation is performed. The outcome of the tool qualification process specifies the confidence level of the software tool usage in developing safety related systems development according to ISO26262.

Until now, we have talked about functional safety, but we have not discussed how the system should respond to functional insufficiencies or in the case of foreseeable misuse by people. These aspects are covered by the SOTIF standard, which complements ISO26262.

One of the objectives of SOTIF is to validate the automated function in all relevant environmental scenarios, especially in difficult conditions for both sensors and algorithms.

For a given operational design domain, SOTIF classifies the scenarios in four areas, as shown in figure 7. The safety goal according to SOTIF is to explore the scenario space during development with an iterative approach, gradually increasing the known safe scenarios area by discovering unknown and unsafe scenarios (area 3) and moving them into known and unsafe scenarios (area 2).

Figure 6. Traffic sign images: normal with pixel errors and with injected fault.

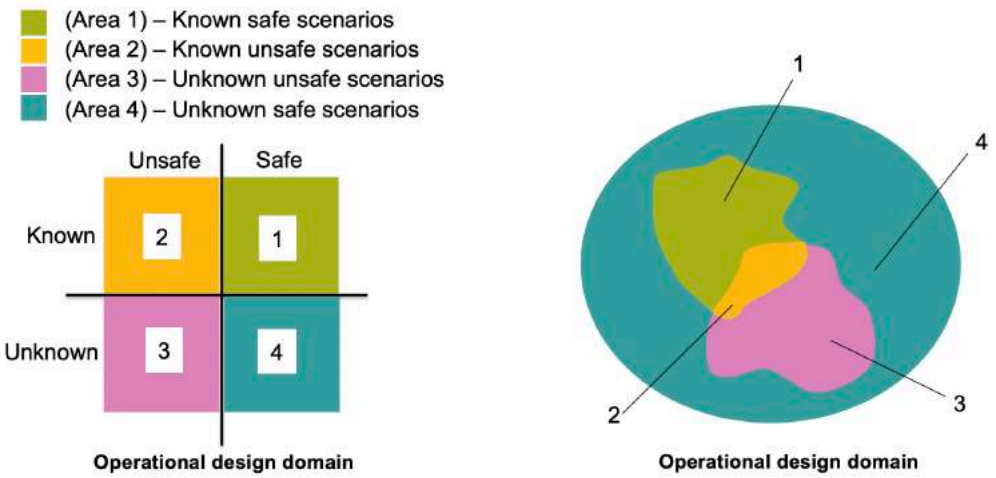


Figure 7. Reducing the risks according to SOTIF.

Then proper technical measures are used to transform them into known and safe scenarios (area 1).

Finally, let us consider the traffic sign detection and classification example from a SOTIF perspective. In this case, scenarios/conditions, which are difficult for both sensors and algorithms, shall be considered. For example, how the detection and classification algorithm will work when the traffic sign is under heavy rain and the camera sensor is aged (see figure 8).

After this brief introduction into two relevant safety standards – related to automated driving systems development – let us make a quick overview of what is state-of-the-art in safety assurance, mainly looking into R&D of different AVs.

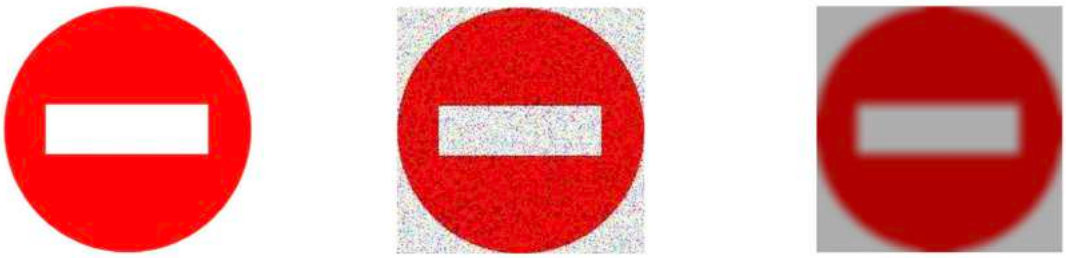

Figure 8. Traffic sign images: normal, under heavy rain and with aged camera lens.

# Safety assurance: state-of-the-art

At this stage, we might ask how safety is assured while developing automated driving systems? If we look at different original equipment manufacturers (OEMs) and new entries in the automotive market developing autonomous vehicles, we observe safety is assured by design, which means: how we define design, develop, deploy an autonomous vehicle/system.

First, let us look at the development of the GM Cruise, in which following proven engineering and development standards eliminated or minimized the risks and two key safety performance thresholds are defined:[5]

- Vehicle will operate safely even if there is a single point, plausible dual point or common-cause malfunction occurs
- Vehicle will demonstrate safe driving behavior in the defined driving environment using a statistically meaningful experience

The first safety performance threshold is achieved as a result of comprehensive risk management and a deep integration process.[4] This process diversifies systems and adds redundancy, which are key drivers of the safety of the Cruise AV, see figure 9.

Furthermore, at GM manufacturing supports system safety, where a "built-in quality" method is used; there are assembly line quality checks for components, subsystems, systems and when vehicle assembly is complete.

The second safety performance threshold is achieved with city testing and proving safe driving with experience. In on-road testing, a fleet of self-driving vehicles that each have a steering wheel, brake pedal and accelerator pedal are used.

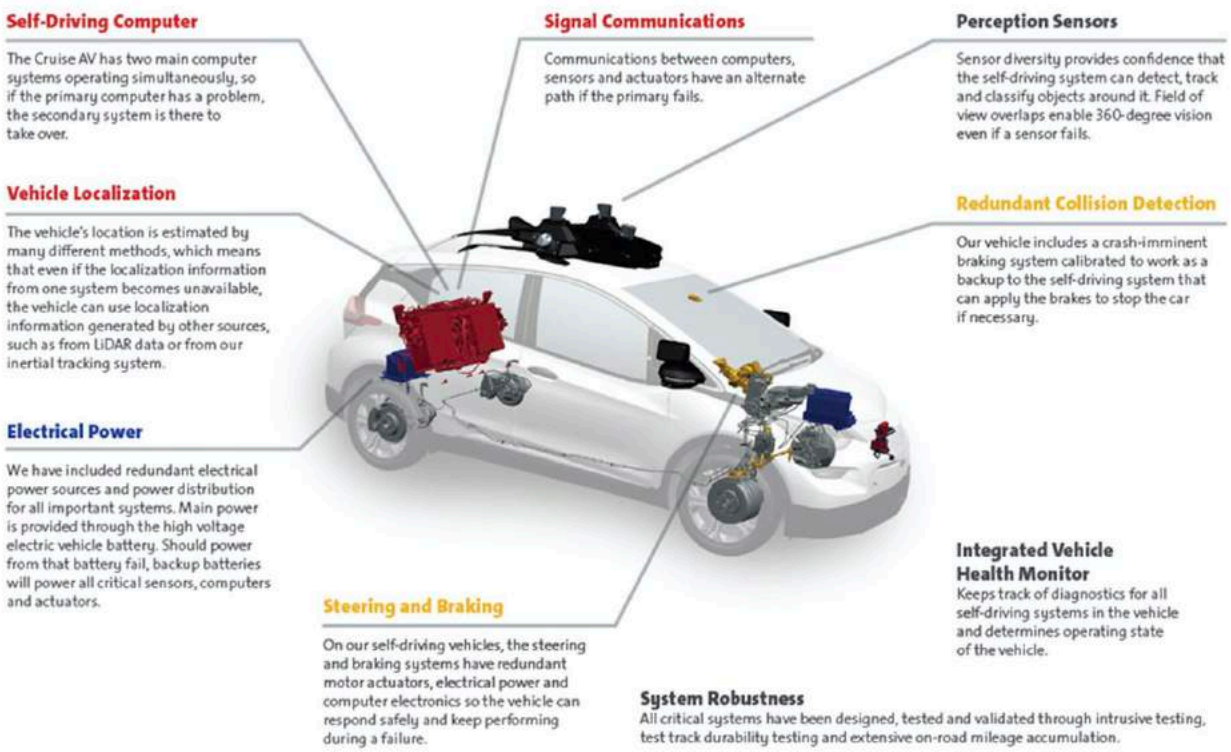


Figure 9. Safety by system diversity and redundancy (source: GM Self-Driving Safety Report 2018).

At GM validation methods include (but are not limited to):[5]

- Simulations for testing the self-driving vehicle against a variety of objective tests and performance requirements
- Track testing and staged encounters
- On-road performance testing, collecting millions of miles of test data to show on a statistically significant basis the vehicle can be driven safely

Validation is viewed as a combination of conventional system validation with SOTIF. Examples of SOTIF validation processes include:[5]

- Systematically expose self-driving system to performance requirements of the operational design domain (ODD)
- Identify and test of driving scenarios and edge cases that challenge the self-driving system
- Exercise the object and event detection and response (OEDR) capabilities of the vehicle and its ability to identify environmental objects and situations that require a safe behavior response
- Evaluate self-driving behavior against safe driving standards with both qualitative and quantitative criteria

The next AV developer we will discuss is Waymo (founded by Google), which has a system safety program governed by the safety-by-design concept.[6] At Waymo, the system safety program (see figure 10) addresses five distinct safety areas:

- Behavioral safety
- Functional safety
- Crash safety
- Operational safety
- Non-collision safety

Each aspect requires a combination of testing methods, which taken together allows Waymo to validate the safety of their fully self-driving vehicles.

Furthermore, Waymo states:[6] "Safety requirements needed to reduce the risk of potential hazards are captured internally, addressed in design, and then verified and validated to demonstrate that safety risks have been reduced to the levels identified in the analyses."


Figure 10. System safety program at Waymo: safety by design.

Waymo's safety requirements include minimal risk condition (fallback), ensuring the vehicle can transition to a safe stop when the self-driving system experiences a problem, is involved in a collision or when the environmental condition changes in such a way that affects the safe driving within the operational design domain.

The software is extensively tested in simulations: The most challenging situations encountered by the vehicles on the public roads are turned into virtual scenarios and virtual verification and validation is performed.

Before real-world driving, closed-course testing is performed on a private test track with the help of experienced drivers. Once it is confirmed the software is working as intended, deployment is initiated to vehicles on public roads. First, the software is deployed on a few vehicles, and after it is confirmed the vehicles can safely and consistently travel on a predetermined route, the software update is performed for the entire fleet.[6]

Scenario-based verification and validation is performed, considering not only the set of behavioral competencies recommended by the National Highway Transportation Safety Administration (NHTSA) (see figure 11), but additional tests are performed on behavioral competencies, such as perform lane change, detect and respond to lead vehicle and detect and respond to a merging vehicle.

In addition, Waymo has developed a robust process to identify, prioritize and mitigate cybersecurity threats. Waymo's security practices are built on the foundation of Google's security processes, the National Highway Traffic Safety Administration's (October 2016) "Cybersecurity best practices for modern vehicles." (Report No. DOT HS 812 333).

Finally, a continuous improvement process is in place at Waymo: Robust data collection and analysis is in place, so anything learned from one vehicle is applied to the entire fleet.

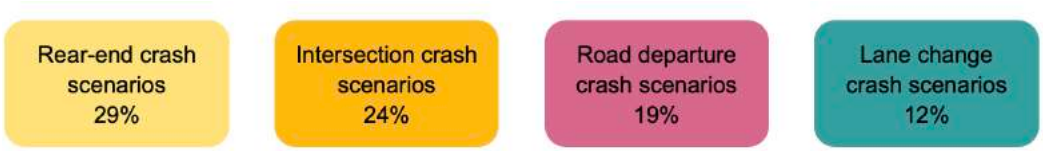


Figure 11. Common crash scenarios according to NHTSA safety report 2018.

# Safety assurance: Siemens Digital Industries Software's perspective

Siemens Digital Industries Software is committed to the safety-by-design concept, on the one hand using the closed-loop development process throughout the life-cycle of the autonomous driving system, and on the other using data continuity throughout the entire supply chain, from chip to city.

The Siemens Digital Industries Software portfolio has simulation and emulation solutions at the following levels: chip, electronics (including the sensors, network and embedded hardware/software), vehicle, fleet of self-driving cars as well as city infrastructure.

With this solution portfolio, the silicon industry develops autonomous vehicle-specific integrated circuits (IC) within the context of the system and vehicle. The automotive suppliers and AV stack technology companies design and verify vehicle-agnostic system implementations using simulation and test software, hardware and engineering services from Siemens.

Finally, the mobility-as-a-service suppliers of the future rely on Siemens Mobility intelligent transportation systems for vehicle-to-everything (V2X) communication and multimodal transportation management.

Furthermore, simulation is heavily used to build functionally safe and secure systems. The breadth of the digital twin portfolio of Siemens, from chip to city, targets continuous integration of mobility solutions for all stakeholders in the supply chain (see figure 12).

Next to the vertical continuity within the supply chain, Siemens Digital Industries Software develops its solution portfolio with an eye toward closed-loop product development cycles, supporting continuous improvement of vehicle performance over the lifetime of the vehicle (see figure 13).

In practice this means simulation models are not only used in the design and exploration stages and for virtual and mixed reality testing, but also during the time the vehicle is deployed. Secondly, Siemens offers solutions for raw sensor data capturing and services for analyzing and diagnosing large data sets, preparing the data for usage in the next round of design and exploration

In the next section let us focus on testing and certification.

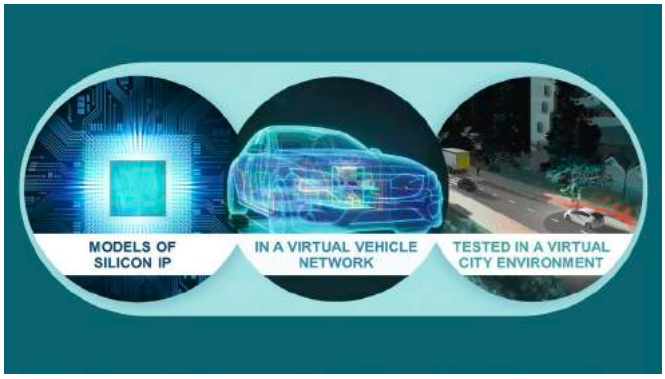


Figure 12. Continuous integration – from chip to city.


Figure 13. Closed-loop development cycle – a continuous improvement process.

# Testing and certification

We have seen that technical and regulatory challenges are strongly linked with business challenges. One of the relevant questions is how to prove that an automated driving system is safe. Since the complexity of systems is growing, in order to test and reproduce extensively critical scenarios (corner cases, edge cases) and keep the costs under control, a significant part of the tests must be done in the virtual world.

Virtual verification and validation will not exclude or limit tests on the test track or in the real-world environment; rather, they complement them.

An obvious question is how much can we rely on virtual verification and validation? Virtual verification and validation, including model-in-the-loop (MiL), software-in-the-loop (SiL), hardware-in-the-loop (HiL), driver-in-the-loop (DiL) and vehicle-in-the-loop (ViL), are based on simulation models. We can rely on simulation models if they are validated against experimental data.[4] Furthermore, the area/region where the model is valid shall be specified and an upper bound of uncertainty of model error shall be calculated and supplied with the model (see figure 14).

Despite their complexity, automated driving systems are composed of three main subsystems: perception, decision and actuation control, as shown in figure 15. When trying to cope with the complexity of the system, we recommend a systematic approach for V&V in which first each subsystem is tested and then integration and testing is performed (see figure 15).

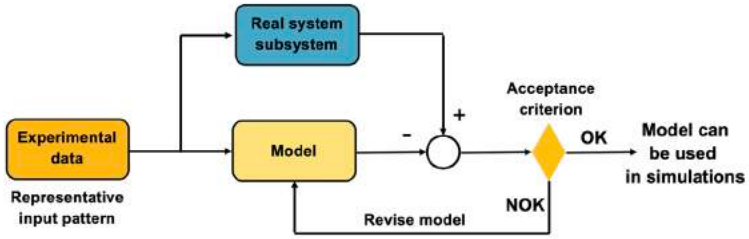


Figure 14. A general framework for model validation.

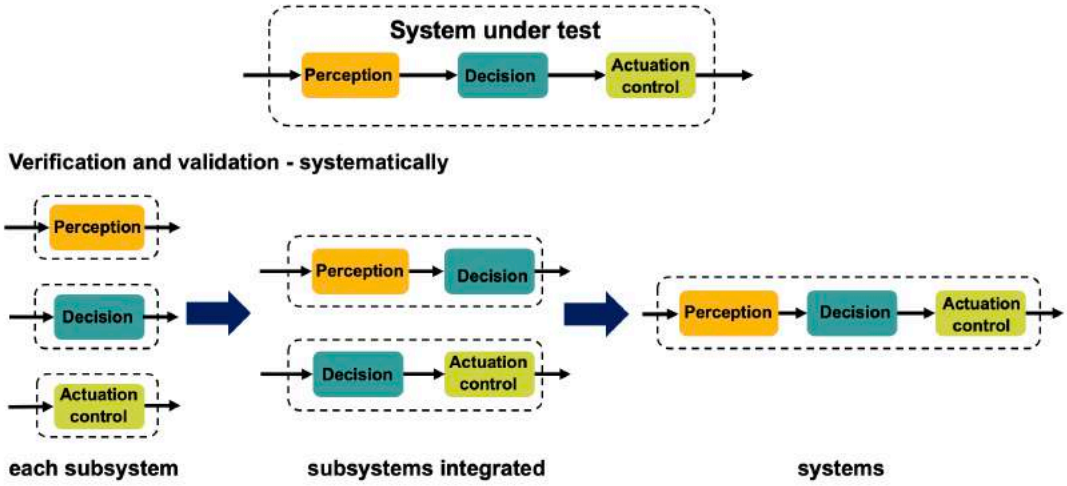


Figure 15. A systematic approach for V&V of an automated driving system.

One of the relevant examples related to test virtualization is the future European New Car Assessment Program (EURO-NCAP) test procedure. EURO-NCAP intends to keep adding complexity (see figure 16) to their test program in order to drive manufacturers to develop the safest cars.[7] This causes problems for the test houses that already have an overloaded test burden for making a single assessment.

Following informal discussions with EURO-NCAP, it is clear they wish to add a level of HiL tests to their program to minimize the burden on labs.

A possible solution is shown on the test grid in table 1. Some of tests are done in the physical world and the data collected are used for model building and model validation. Most of the rest of the tests can be done in the virtual world, such as HiL tests. This approach might avoid significant workload increase at test houses due to the increase in test complexity.

Siemens Digital Industries Software in The Netherlands has its own test facilities to drive with autonomous vehicles in mixed traffic, including 5G-connected vehicle testing.

Here test facilities include a highway test environment, urban and interurban test environment, industrial area test environment, 5G test facilities and computer-aided design (CAD) car labs.

Siemens' combination of software tools and test facilities offer excellent solutions and services that cover the entire closed-loop development of automated driving systems based on a digital twin.

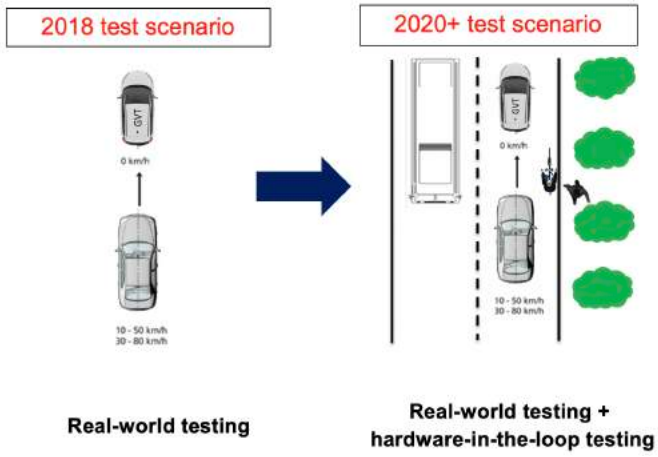


Figure 16. Complexity increase in case of future EURO-NCAP tests.

| | | AEB interurban | | | | |
|---|---|---|---|---|---|---|
| | | Lateral overlap [%] | | | | |
| | | -50% | -75% | 100% | 75% | 50% |
| EGO velocity | 30 [km/h] | Measured | • | • | • | Measured |
| | 35 [km/h] | • | • | Measured | • | • |
| | 40 [km/h] | • | • | Measured | • | Measured |
| | 45 [km/h] | • | • | Measured | • | • |
| | 50 [km/h] | Measured | • | • | • | Measured |
| | 55 [km/h] | • | • | • | Measured | • |
| | 60 [km/h] | • | Measured | Measured | • | • |
| | 65 [km/h] | • | • | • | • | • |
| | 70 [km/h] | • | • | Measured | • | • |
| | 75 [km/h] | Measured | Measured | • | • | • |
| | 80 [km/h] | • | • | Measured | • | • |

| Legend | |
|---|---|
| **Measured data** | |
| Build model | (fitting the model) |
| Validate model | (cross validation) |
| **Prediction** | |
| • | (estimated value) |

Table 1: Combining real and virtual testing in case of future EURO-NCAP tests.

A good example is the ITS Europe 2019 demonstration (in cooperation with partners) for design, exploration and validation of connected mobility systems using an intelligent intersection and vehicle-to-vehicle (V2V) communication (see figure 17, left-side digital twin, right-side real-world demonstration).

Referring to testing of autonomous vehicles, for an objective comparison of the test results general safety metrics have been defined. Below are three mathematical models related to autonomous vehicle safety:

- Mobileye's Responsibility-Sensitive Safety[8]
- NVIDIA's Safety Force Field[9]
- IVEX's, IVEX Safety Assessment

Although we have mainly discussed ISO26262 and SOTIF in this white paper, we would like to highlight that existing standards do not present guidance for some of the most problematic topics of automated driving systems:

- Safety assurance of artificial intelligence
- Technological capabilities of sensory devices
- Human factors and psychology.

For example, the well-established standard ISO26262 talks about fail-safe behavior. However, in case of automated driving systems the new design should create fail-operational or fail-degraded behavior. Furthermore, the emerging standard SOTIF looks only at L1 and L2 automation (a revision is underway to address higher automation levels).

We can conclude that there is no unified standard, which might guide the certification of autonomous vehicles for example, level 4 automation). Despite this there are significant efforts to create such region-specific standards, such as Technical Reference 68 (TR68) from Singapore.

The Safety Case Framework Report 2.0 (by ZENZIC from the United Kingdom) summarizes the safety cases for testing and developing connected and self-driving technologies. The safety case must demonstrate the field tests and real-world tests are being conducted in a safe manner and in accordance with United Kingdom law.[10]

For test facilities in the public domain, it is required to comply with United Kingdom road traffic laws as well as the requirements of the land owners and good practice. Evidence of required compliance should be provided in the form of an assessment and declaration/statement of compliance with the relevant standards and regulations.

A safety case level matrix – combining the confidence level for the three main factors below – will provide the assessors with the level of safety case required:

- Safety operator confidence levels and considerations assessing confidence
- Vehicle confidence levels and considerations assessing confidence
- Environmental control confidence levels and considerations assessing confidence

The matrix is a qualitative tool that can guide testbeds and test organizations but should allow for the flexibility to exercise professional judgement.[10]


Figure 17. Scenario simulation using V2X and real-world demonstration at ITS Europe 2019.

# Conclusion

In this white paper we briefly review the relevant standards currently used in ensuring safety for automated driving systems, ISO26262 and SOTIF. We present the main safety assurance measures, which were taken during development and deployment of the GM Cruise and Waymo (Google). We discuss trends in testing and certification in light of the Euro-NCAP 2025 roadmap as well as from the perspective of autonomous vehicles targeting automation level 4.

Siemens Digital Industries Software has been on a well-defined path for years to support the automotive industry in the transition toward connected, autonomous, shared and electric mobility. The unique combination of our simulation software and testing services, stretching from chip to city, and the portfolio evolution towards seamless definition, design, development and deployment, aims to support the industry to meet the safety expectations of the coming decade.

## References

1. Safety First for Automated Driving, 2019.
2. Road vehicles - Functional Safety, ISO 26262, second edition, December 2018.
3. Road vehicles - Safety of the intended functionality, ISO/PAS 21448, first edition, January 2019.
4. Forrai, Alexandru: Embedded Control System Design: A Model Based Approach, Springer 2013.
5. General Motors: SELF-DRIVING SAFETY REPORT, 2018.
6. Waymo Safety Report: On the Road to Fully Self-Driving, 2018.
7. Euro NCAP 2025 Roadmap, 2019.
8. Shalev-Shwartz, Shai; Shammah, Shaked; Shashua, Amnon. *On a Formal Model of Safe and Scalable Self-driving Cars*, Mobileye 2017.
9. Nister, David; Hon-Leung Lee, Hon-Leung; Julia Ng, Julia; Yizhou Wang, Yizhou. *The Safety Force Field*, NVIDIA, March 2019.
10. Zenzic: *Safety Case Framework Report 2.0*, March 2020.

**Siemens Digital Industries Software**

**Headquarters**
Granite Park One
5800 Granite Parkway
Suite 600
Plano, TX 75024
USA
+1 972 987 3000

**Americas**
Granite Park One
5800 Granite Parkway
Suite 600
Plano, TX 75024
USA
+1 314 264 8499

**Europe**
Stephenson House
Sir William Siemens Square
Frimley, Camberley
Surrey, GU16 8QD
+44 (0) 1276 413200

**Asia-Pacific**
Unit 901-902, 9/F
Tower B, Manulife Financial Centre
223-231 Wai Yip Street, Kwun Tong
Kowloon, Hong Kong
+852 2230 3333

## About Siemens Digital Industries Software

Siemens Digital Industries Software is driving transformation to enable a digital enterprise where engineering, manufacturing and electronics design meet tomorrow. Our solutions help companies of all sizes create and leverage digital twins that provide organizations with new insights, opportunities and levels of automation to drive innovation. For more information on Siemens Digital Industries Software products and services, visit siemens.com/software or follow us on LinkedIn, Twitter, Facebook and Instagram. Siemens Digital Industries Software – Where today meets tomorrow.

**siemens.com/software**



81936-C4 4/20 A