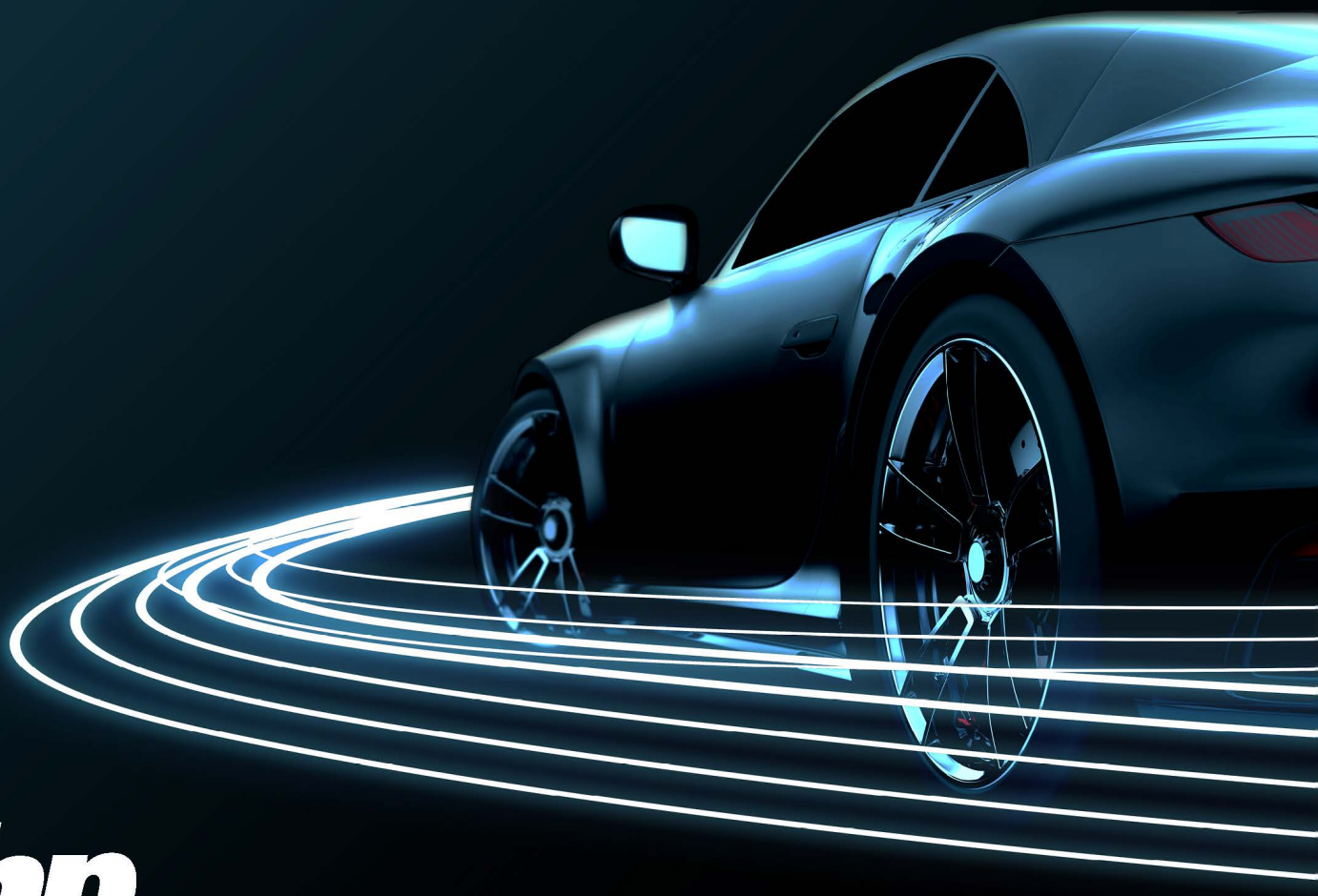
eBook

THE GUIDE TO SOTIF ISO 21448

The Key Driver of Safety for Autonomous Vehicles



LHP®

Table of Contents

About the Publisher 7
Foundation concepts 8

A quest to build a safer world 10
Taming the complex 10
Defining “safe” 10
The evolution of the modern mechatronic vehicle system 11
A reactive past 11
Deliberate safety intent 12
The limitations of human-based data processing 13
A standard to guide you 13
How does the SOTIF standard help to ensure functional safety? 14
Technical competence 14
Designing the scope with deliberate intent 14
Managing the risks that drive the work 15
The risk of human error 15
Defining the risks 16
Defining an acceptable level of safety 16
Advanced Driver Assistance Systems (ADAS) 17
What is an Advanced Driver Assistance System? 17
A history of safety in ADAS 17
A collaborative relationship 18

An overview of ISO 21448, the SOTIF standard 19
The scope of ISO 21448 19
What is in the scope of ISO 21448? 19
What is out of the scope of ISO 21448? 20
Ensuring the safety of the intended functionality 21
Achieving and maintaining the SOTIF 22
Applying a vetted methodology 22
Situational awareness and autonomy 23
Managing misuse 24
The impact of proximity 24
The References 25
Safety-relevant topics addressed by other standards 25
Normative references in the standard 26
Examples of references 26
Terms, definitions, and external sources 27

The SOTIF principles 28
The SOTIF-related hazardous event model 28
Proper situational awareness 30
Changes to situational awareness 31
Situational variations 31
Managing changes to situational awareness 32
Changes that impact the driving policy 32

The SOTIF scenarios 33
Evaluations and probabilities 36
Unknown hazards versus known hazards 37
Unknown hazards 37
Known hazards 37
Evolution and progress 38

The Sense-Plan-Act model 40
How does the Sense-Plan-Act model work? 40
A key part of the system architecture 41

How ISO 21448 is used to achieve the SOTIF 42
The detailed order of SOTIF activities 42
The flow of SOTIF activities 44
The Annexes in ISO 21448 45

The management of the SOTIF activities and their supporting processes 47
Parallel activities between the standards 47
SOTIF development activities for distributed product development 47
SOTIF-related elements and their relationship to context 48

The specification and design of the SOTIF 49

Considerations for the design of the system and its architecture 51
The importance of an all-encompassing description of the system 51
Achieving understanding 51
Continuous improvement 51
Methodically encompassing previous work in new updates 52
Cooperation among the development parties 52
Traceability 52
Performance insufficiencies and their countermeasures 53
The identification and evaluation of hazards 54
Sources of hazard information 54
Identifying the hazards 55

The evaluation of risk 56
The severity and controllability evaluation 56
The evaluation of delayed reactions or the lack of reactions by people 56
Specification of the acceptance criteria for residual risks 57

Identifying and evaluating potential functional insufficiencies and potential triggering conditions 58
Combining methods 58
Addressing multiple triggering conditions 58
Analysis of reasonably foreseeable misuse, either direct or indirect 59
Estimating the acceptability of the system's response to triggering conditions 59

Managing the functional modifications that address SOTIF-related risks 60
Considerations for improving the SOTIF 60
Refining the system 60
Modifying the system 61
Functional restrictions 61
Handing over authority from the vehicle to the driver 62
Effective strategies for addressing reasonably foreseeable misuse 62

The verification and validation strategy 63
The scope and purpose of verification and validation activity 63
Integration and testing 64

Evaluating the known and unknown scenarios 65
Evaluating known scenarios 65
Evaluating unknown scenarios 65
Testing in public areas 66
Length and method of testing 66

Evaluating the achievement of the SOTIF 67
Methods and criteria for evaluating the SOTIF 67
Recommendation for the release of the SOTIF 69

Operation phase activities 70

Unsolved challenges 71
The challenge of managing updates 71
Retroactive compatibility 72
Updating from lessons learned 72
The importance of accurate and complete data 73

Summary 74
Bibliography 75

ABOUT THE PUBLISHER

Thank you for downloading the eBook, “The Guide to SOTIF ISO 21448.” This latest eBook should be used as a resource to further educate your organization on the paramount importance that the Safety Of The Intended Functionality (SOTIF) plays in the functional safety of products and their systems.

LHP Engineering Solutions is a technology integrator and engineering services provider in the transportation industry. LHP provides expertise to the automotive industry on topics including ADAS, ASPICE, SOTIF, AUTOSAR, Cybersecurity, and ISO 26262.

Our mission is to create a safer, smarter, and more connected world using autonomous technology and their supporting frameworks. We believe that by developing technology solutions and engineering services built on intelligence, connectivity, and actionable insight, real autonomy is possible.

We hope you enjoy this eBook and are one step closer to a safer autonomous vehicle.

Thanks,

David Glass, CEO

The Guide to SOTIF ISO 21448: The Key Driver of Safety for Autonomous Vehicles

Fast-forward to the not-so-distant future. There it is, a gleaming new modern functionally safe vehicle in your driveway, complete with that new car smell. How did it get there? How did this functionally safe vehicle come into existence? How did the designers and builders confirm that this vehicle is indeed functionally safe?

Let's start at the ***end*** of this journey and ***work backwards:***

- In order to have a **functionally safe ecosystem**, the vehicles and systems operating within the ecosystem must first be proven to be functionally safe.
- In order to certify that **a given vehicle is functionally safe**, the vehicle and all its systems must first be thoroughly tested. At a latter point in the process this testing will involve road testing in the real world, but teams don't start there.
- In order to test all the systems that comprise the vehicle and the vehicle itself as a whole unit, there must first be a logical, consistent, and repeatable development and testing process to follow. The **Veri ication and Validation (V&V)** process breaks high-level complex challenges down into finer and more basic levels step-by-step, and then builds them back up in complexity to the point that real-world testing becomes viable. This en-sures that things are being measured in a trustworthy and repeatable way, so that the re-sults of the tests are consistent and legitimate, and can be acted upon with confidence.

- In order to conduct verification and validation, there must first be **clear and unambiguous test criteria** to aim for and measure against.
- In order to define the test criteria, **the scope of each test must first be defined** so that testers don't find themselves trying to boil the ocean. These scopes become clarified by:
    - **Measuring and confirming the tangible things** that you can touch, create, manipulate, and interact with, such as:
        - the management of the systems engineering;
        - concepts of operations;
        - system level requirements;
        - sub-system requirements;
        - component detailed designs; and
        - implementation hardware and software coding and testing.
    - **Defining the scenarios** that detail the mix of hardware, software, environmental conditions and other elements impacting a given place and situation at a given point in time. In other words, accurately defining the world in which the vehicle will operate.

To conduct all this work effectively, many different processes and methods are followed, but there are shared common themes:

- Start at a high level, and then break items and challenges down into their most basic elements until they can't be realistically broken down any further, and first solve problems at this fundamental base level.
- Like building a toy house with blocks, build complexity one step at a time, only after you have confirmed that the elemental blocks you are building upon are accurate and complete. Don't try to build on vague or unproven data.
- Conduct all your work according to vetted standards and the processes they define. No shortcuts.

All of these processes and themes describe both the methodology and spirit applied to achieving SOTIF, the Safety Of The Intended Functionality. The standard that governs SOTIF activities is ISO 21448:2022, *Road vehicles — Safety of the intended Functionality* [1]. But how did the concept of SOTIF, and its standard, come to exist? What motivators drove their development and refinement? And how is the standard used today to guide and drive functional safety?

A quest to build a safer world

Taming the complex

The operation of a modern vehicle involves a mix of automatic, semi-automatic, and human-directed systems. This assemblage is quite complex. But adding exponentially to this complexity, is the reality that this vehicle must share a widely varying network of roads, in ever-changing conditions, with other vehicles that also have different and varying blends of systems and capabilities. And, they are all being driven by unique individuals possessing widely-varying sets of skills, priorities, experiences, and abilities, whose levels of concentration vary constantly.

This infinite variation creates an incredibly complex and ever-changing environment in which the modern mechatronic vehicle must operate.

Despite this broad and dynamic variation, there is a point of commonality among all these vehicles and systems, the one high-priority consideration that binds them together and drives their designers towards a common goal: safety. After all, our vehicles exist to serve humans and make our world a better place. If they can't be used without them hurting us in the process, what is the point of having them?

Defining "safe"

In this realm, the word "safe" has specific meaning. Either something is safe, or it is not. Yes, a vehicle or system can be deemed safer or less safe than it was before, or, safer or less safe than something else. But if the vehicle is deemed to be fully safe in the presence of reasonable risk, it is considered *functionally safe*. This is a formal term that is challenging to achieve and must be earned through rigorous process work and evaluation.

If a vehicle is functionally safe, that means that it remains in a safe state regardless of whether it is being operated as its designers intended, or it is being operated under conditions or in a manner unforeseen by its designers.

In this eBook, we will examine the SOTIF processes and standards that govern SOTIF work. However, to fully understand SOTIF, we must also review a bit of automotive history to better understand how SOTIF came to fit into and interact with the other elements and considerations that impact functional safety today.

The evolution of the modern mechatronic vehicle system

A reactive past

The vehicles of today are a far cry from their latter 20th Century predecessors. Long gone are the days when electrical and mechanical systems were essentially separate systems, augmented by rudimentary control units bolted to the firewall.

Today's modern vehicles are highly complex mechatronic machines that tightly interweave mechanical systems, software, computational power, sensing, data and bandwidth capacity, and physical actuation, implemented by electrical and electronic systems.

The interwoven nature of these systems stretches from bumper to bumper and beyond, reaching out to include the environment the vehicle is operating in. This capability enables the mechatronic system as a unified whole to make decisions and take actions based on what it perceives, and what it has been designed to do.

It can be said that today's safer vehicles and safety efforts can be traced back to the earliest automotive safety efforts. When automobiles were first introduced, they were little more than a stepped progression from the horse-drawn wagons and carriages they replaced. Neither the vehicles themselves nor the roads they operated on were designed to accommodate the speeds, weights, and quantity of vehicles that the rapid innovation of the automobile would introduce.

The early automotive pioneers designed their vehicles primarily through experimentation and reaction. If something worked, they simply tried to replicate the success even before they fully understood *why* it worked. And just as lessons were learned from what worked, poignant lessons were also learned by what didn't work, sometimes at a tremendous cost in human suffering.

Deliberate safety intent

Slowly, the process started to move away from being purely reactive to having deliberate safety intent designed into it, reflecting the earliest hints of what would eventually become SOTIF, the key word in the acronym being *intention*.

The design of vehicles and roads started to transition from being reactive, to having safer functionality that was the product of intentional forethought. Sound engineering principles were applied to both the design and construction phases, improving the stability and safe carrying capacity of both the vehicles and the roads and bridges they operated on.

Engineers combined scientific studies with environmental and materials science to introduce safer vehicles. America began building better roads. Correspondingly, vehicles were fitted with better tires specifically engineered for these new surfaces, improving safety by providing a better grip and a more stable ride. The roads themselves evolved into engineered systems built to strict and consistent universal standards that were designed and exhaustively tested under the guidance of various highway-building associations and, eventually, federal regulatory agencies.

There are key take-aways from all this history:

- Safety through deliberate forethought is a concept that predates the automobile and is more important than ever today.
- We define and apply intended functionality, because we have sound data to prove that doing so makes our world safer.
- Applying vetted and repeatable scientific principles is more effective than guesswork.

All of this history and experience underscores the importance of designing to a deliberate safety intent.

The limitations of human-based data processing

As time went by, data *accuracy* improved with the increased sensitivity and ruggedness of improved instrumentation, but the data were still being *processed* in real-time, by humans. This data was the product of imperfect processing and subjective storage in the form of scribbled notes and fallible human memory. It was limited as much by the bandwidth constraints of the human brain as by the technology.

The shortcomings in these systems revealed the need to improve them beyond the capabilities of imperfect human idiosyncrasies. Better sensors and instrumentation were developed to transcend the humans in order to better serve them.

Today, the typical modern vehicle is equipped with sensors, actuators, and computational power thousands of times more advanced than those that helped the space program put men on the moon. The quantity and complexity of the components and data now being utilized, and the ability to utilize them with greater effectiveness, has necessitated the creation of safety standards like SOTIF to, among other considerations, manage the complexity of it all.

A standard to guide you

Before you can formally define the safety of the intended functionality, you must first have a standard to provide a defined, measurable, and repeatable process for achieving the safety goals. ISO 21448:2022, *Road vehicles — Safety of the intended functionality* is that standard.

Recently updated in 2022, ISO 21448 is an ISO standards document created and maintained by the International Organization for Standardization (ISO), a worldwide federation made up of national-level standards bodies. Robust and properly vetted, this standard helps ensure the proper implementation of functional safety.

How does the SOTIF standard help to ensure functional safety?

Technical competence

As noted in the standard itself, the preparation of International Standards is normally carried out through ISO technical committees. International organizations (both governmental and non-governmental) in liaison with ISO, also help support this work. Through the diligent work of these entities and persons, the standard is continuously evaluated and improved.

In order, the process journey by which ISO 21448 is prepared and evolves through various drafts and reviews, and is eventually approved for publication, is as follows:

1. Subcommittee SC 32, Electrical and electronic components and general system aspects;
2. Technical Committee ISO/TC 22, Road vehicles;
3. ISO approval and publication.

Note that the SOTIF standard originates in a subcommittee from the realm of electronics and general systems, not at the road vehicle level. This makes practical sense, given that today's modern vehicles are electromechanical, and some of those same sensors and components might be shared with other types of vehicles and applications including on-highway, off-highway, or even shared with realms outside automotive such as aerospace applications.

It is important to grasp this added complexity right from the start. Long gone are the days when a vehicle was simple enough to be subdivided into separate electrical and mechanical systems. Each modern application necessitates defining requirements that have been tailored to that particular use case and operating environment.

Designing the scope with deliberate intent

The first and most basic principle of understanding the scope of SOTIF, is realizing that its scope is not defined by malfunctions, but rather, the *effectiveness* of the intended functionality. The scope of SOTIF defines properly working equipment that has been designed and built and tested to fulfill their given requirements. **At its most basic, SOTIF is scoped to specific intentions manifested through thoughtful and deliberate engineering, and the application of lessons learned.**

As teams work through the SOTIF process, they don't constantly move the borders of the *scope* as new information comes in, reacting haphazardly to every new and unexpected return in a patchwork of duct tape and bandages. Instead, **they try to figure out whether the vetted and approved requirements for a given system are good enough for what it was *designed* to manage.**

In other words, the SOTIF acronym might be expanded to be thought of as safety of the intended functionality of a **particular system.** Were the requirements of that *particular system* properly specified? Did we capture the intent of that *particular system?* The team works its way down the list and asks the same questions, over and over, for each safety-related system.

Managing the risks that drive the work

In recent years, there has been a large increase in the number, capability, and complexity of advanced vehicle functions. Bit by bit, the capability of the vehicle is growing to the point that it can augment, or in some cases take over, functions that previously could only be handled by humans.

However, with increased complexity also comes increased risk. To efficiently manage these risks, it makes sense for functional safety efforts to focus on the areas of greatest risk, and then try to reduce or eliminate them.

The risk of human error

Because most vehicle functions are controlled by humans, most vehicle accidents can be traced to human error. Therefore, the most effective way to reduce the number of accidents, is to provide trustworthy assistance to the human that controls the vehicle.

This is accomplished by off-loading select decisions and capability to automated systems that have been proven to perform these specific tasks faster, more accurately, and more reliably, than the typical human.

The intent is not to replace the human per se, but for the human and the vehicle to work together to free the human to focus on those tasks that the humans most wish to perform.

Defining the risks

Before any design work can be completed, one must define the risks. After all, you must quantify the risks that you are trying to mitigate, before you can design and build systems to mitigate them.

Risks are identified and sorted into one of four categories:

- Known not hazardous
- Unknown not hazardous
- Known hazardous
- Unknown hazardous

Because each type of risk carries its own idiosyncrasies, each category is prioritized and dealt with in a slightly different way.

Defining an acceptable level of safety

Defining the scope of SOTIF involves lots of steps, each with a purpose, and they must be performed in the proper order. They must address both the intended functionality of the system, and any unintended functionality that may arise.

ISO 21448 details how automotive systems should be verified and validated as being functionally safe. This must be accomplished for every system on the vehicle that can impact safety in any way.

For a road vehicle to achieve an acceptable level of safety, unreasonable risk must be avoided. Not diluted, not minimized, but *avoided entirely*. This includes every hazard that is associated with both the intended functionality of the vehicles, and any of the unintended functionality resulting from its use in the real world.

Advanced Driver Assistance Systems (ADAS)

What is an Advanced Driver Assistance System?

To better understand SOTIF, it is helpful to understand the tangible end products created through the SOTIF process, and other related processes. Advanced Driver Assistance Systems (ADAS) are the vehicle hardware, software, and communication systems that go into a functionally safe vehicle. They are designed to increase safety and reduce risk by using the human-machine interface to aid the humans driving the vehicle.

ADAS systems provide early warnings and reduced reaction times to potential hazards. This enhanced capability helps prevent deaths and injuries by lowering the number of vehicle accidents, and by reducing the serious impact of accidents that cannot be avoided.

A history of safety in ADAS

There are many examples of ADAS systems that have been in common use for decades, including anti-lock braking systems that were first introduced in the 1970's, traction control, headlights that automatically turn themselves on in low light conditions, and rearview mirrors that automatically dim. More recent innovations include rear-view cameras to help prevent backing accidents, adaptive cruise control that allows you to set and maintain a fixed distance from the car in front of you, navigation assistance in the form of GPS-based graphical and audible systems, hazard avoidance around the vehicle, and lane departure and centering.

The technologies used in these systems can typically be categorized as either those that **improve driver awareness**, or those that **automate driving tasks**.

These technologies were selected, and these systems were designed and built, only after engineers first defined:

- what safety problems the systems were intended to solve;
- what a "safe" system looked like; and
- what known and unknown risks might be encountered in their use.

Defining these criteria, is the act of defining the safety of the intended functionality. But how do these systems come into being? This is where SOTIF processes come into play.

1. ADAS systems must be designed, built, verified, and validated. This work is performed by a variety of teams in a controlled development environment, before the vehicle is deployed into the real world.

2. Safety criteria must be established before first design, so the designers and builders know what to strive for. These same criteria are used to measure whether the completed system achieves the safety goals.

3. Once the system is deployed out in the real world, the intended functionality is measured against the actual functionality that the system achieves when faced with both known hazards, and those hazards that were not anticipated.

4. Lessons learned from both the intended and unintended functionality, are rolled back into the safety criteria in the form of improvements. The improvements go through the same process in the development environment, and after verification and validation, are rolled out in a controlled manner at the next production update of the system.

And, guiding all this work are the key international standards that detail how to perform the work, and how to measure its effectiveness.

A collaborative relationship

Overarching the scope of SOTIF, is the realization that it does not exist and function in isolation. SOTIF is a process closely linked to other standards and their processes, such as ISO 26262. The product that is produced as the output of properly applying all these standard processes is the Advanced Driver Assistance Systems (ADAS) that are engineered into the vehicle and that the drivers are protected by and interact with.

Defining precisely what the SOTIF is, establishes the criteria that need to be met when an ADAS system is designed, built, and deployed. In other words, SOTIF defines the target, what the intended definition of “functionally safe” is in each scenario, and ADAS systems are how you hit that target, by turning that intended safety into actual safety in the real world.

If teams don’t first define the intended functionality via SOTIF, they have no way of knowing what they are aiming for, and no way of confirming they achieved their goals. And if they don’t follow through by manufacturing and deploying vetted ADAS solutions, real-world safety is not improved.

An overview of ISO 21448, the SOTIF standard

Note: Throughout this document, we describe some of the highlights of ISO 21448 at a more informal and conversational level than the standard itself, but this guide is not a substitute for the standard. All inferences to the standard within this entire document are made under the blanket citation [1] noted in the Bibliography, ISO 21448:2022, *Road vehicles — Safety of the intended functionality.*

Note: Throughout ISO 21448, some words follow European norms for spelling and grammar, while in this eBook, U.S. spelling and grammar are used. EXAMPLE: “visualisation” versus “visualization”. When citing a specific section in ISO 21448, the citation reflects the spelling used in the standard. In the conversational content of this document, the U.S. spelling is used.

The scope of ISO 21448

To better understand the SOTIF standard and how it supports SOTIF activities, it is helpful to review what is and isn’t covered in the standard, as well as to examine some of the areas where the SOTIF standard works closely with other related standards.

What is in the scope of ISO 21448?

The scope of ISO 21448 includes:

- a framework of measures to ensure the safety of the intended functionality;
- guidance on achieving and maintaining the SOTIF;
- guidance for defining the intended functionalities; and
- methods for defining and addressing misuse.

ISO 21448 is not intended to address every conceivable type of vehicle that might be found out on the road, because such a “boil the ocean” standard would be so unwieldly as to be impractical. Instead, the standard is limited in scope to series production road vehicles, excluding mopeds, that utilize one or more electrical and/or electronic (E/E) systems installed at the time that the vehicle was manufactured.

What is out of the scope of ISO 21448?

ISO 21448 does not apply to:

- Faults that are addressed in the ISO 26262 series. (SOTIF is a subset discipline that falls under the ISO 26262 umbrella, not the other way around.)
- Threats from the realm of cybersecurity, which is such a broad and detailed ever-changing topic, that it has its own family of standards and procedures.
- Hazards that can be caused by the technology of the system (example: harm to the eyes that can be caused by a lidar beam).
- Hazards caused by or related to smoke, fire, heat, electric shock, radiation, flammability, reactivity, toxicity, the release of energy, and other similar hazards, unless the hazard is directly caused by the intended functionality designed into the E/E systems.
- Deliberate actions that are in clear violation of the intended use of the system; this is considered feature abuse. EXAMPLE: Modifying the system that governs maintaining a safe distance to the vehicle directly in front, to automatically enable extremely close and hazardous racing-styled “drafting”.
- The functions of systems that already exist, for which vetted design V&V measures are already in place. EXAMPLES: dynamic stability control systems, airbags, etc.
- Aftermarket add-ons, modifications, or the use of parts or maintenance procedures not approved by the original equipment manufacturer (OEM).

Ensuring the safety of the intended functionality

ISO 21448 provides a framework and guidance for ensuring the safety of the intended functionality, defined in ISO 21448 as the absence of unreasonable risk due to a hazard caused by functional insufficiencies[1], such as:

- Insufficiencies in the way the intended functionality was specified at the vehicle level; or
- Insufficiencies in either the specifications, or the resulting performance, when electric and/or electronic (E/E) elements in the vehicle system are implemented.

The manner in which the function of the vehicle is designed and specified, must adequately reflect a truly safe state when the vehicle encounters the conditions found in the real world.

The following two distinctions are important:

- Fulfilling the requirements of the standard helps to ensure that the intended functionality is accurately and completely specified. After all, you must first define a target before you can determine if you have hit it.
- Examining insufficiencies in either the specifications or the actual performance of elements in the vehicle system, differentiates the examination of the systems and subsystems, from the examination of the overall vehicle as a whole.

It is theoretically possible for a vehicle to be doing what it was designed to do, only to learn later that the intended functions designed into the vehicle were misaligned with what was actually needed. For example, a braking system designed for the urban SUV market might be designed to perform well at high speeds on paved roads, only to find out later when an owner hooks up a camper and takes it on vacation, that this same braking system is inadequate for hard sustained downhill use under heavy loads on umpaved roads in mountainous terrain.

Given that the whole purpose of designing and building these vehicles is to use them in the real world, there should be no gaps between intended functionality and real-world needs.

Likewise, it is also theoretically possible for a vehicle to be designed in such a way that a deficiency in the specification or performance of one system might be inadvertently compensated for or otherwise masked by other systems, thus having the undesirable effect of

obscuring the root cause of an issue. For example, a slow loss of fuel pressure to an injector might be compensated for by the fuel system supplying additional fuel to the entire system, but it doesn't address the root problem that there could be a high-pressure fuel leak, which is potentially a very hazardous situation.

Or in another example, an electronic component inside a part might be rated to withstand a certain amount of heat and vibration for intermittent periods of short duration, but not be able to withstand the actual conditions encountered under sustained heavy-duty use once the part is installed in a commercial vehicle that is kept in almost constant operation.

Whether at the vehicle level or subsystem level, all systems should be accurately designed to the true real-world conditions, and they should perform in the real world as intended.

Achieving and maintaining the SOTIF

Applying a vetted methodology

ISO 21448 provides guidance on measures and activities that are needed to achieve and maintain the SOTIF, including:

- Design measures
- Verification and validation (V&V) measures
- Operation phase activities

Having the standard define the proper measures and activities, takes the process guesswork out of achieving the SOTIF and enables the consistency required to produce accurate and trustworthy results that can be understood by partner engineering teams and then acted upon with confidence.

By following the standard, the OEM teams don't have to develop processes from scratch, they only have to implement the existing vetted processes detailed in the standard. This makes it economically viable for a manufacturer to achieve and maintain the requirements of the SOTIF without becoming distracted by the onerous task of creating and maintaining their own unique safety methodologies.

Situational awareness and autonomy

The vehicles that are encompassed by ISO 21448 all share the trait of having functionalities where proper situational awareness is essential to safety. But situational *awareness* is not the same thing as *autonomy*. The two terms, while related, mean two very different things, and are too often utilized inappropriately in common usage as if they were interchangeable or somehow inextricably linked.

A vehicle possessing some degree of situational awareness, is nothing new. The earliest vehicles gained a crude degree of situational awareness the moment they were fitted with the first temperature gauges and speedometers. But it still fell on the human driver to process the information and act appropriately upon it.

In modern vehicles, situational awareness can be found in non-autonomous, semi-autonomous, and autonomous vehicles, although we have just about reached the point where all vehicles of recent manufacture possess some degree of autonomy, even if it is nothing more than automatic braking systems (ABS) or a rearview mirror that dims automatically. In many instances, yesterday's options have become today's standard equipment.

This situational awareness is achieved via complex sensors and processing algorithms, especially in the functionalities of emergency intervention systems, and systems having levels of automation from 1 to 5, as defined in *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles,* SAE Recommended Practice J3016_201806, https://www.sae.org/standards/content/j3016_201806 [2].

Derived from ISO 21448:2022, Table 2 — Levels of Driving Automation[1]

		Dynamic Driving Task (DDT)			
Level	Name	Lateral and longitudinal vehicle motion control	Object and Event Detection and Response (OEDR)	DDT fallback	Operational Design Domain (ODD)
0	No driving automation	Driver	Driver	Driver	Not applicable
1	Driver assistance	Driver and system	Driver	Driver	Limited
2	Partial driving automation	System	Driver	Driver	Limited
3	Conditional driving automation	System	System	Fallback-ready user	Limited
4	High driving automation	System	System	System	Limited
5	Full driving automation	System	System	System	Unlimited

Lane guidance, cruise controls that maintain a prescribed distance from the vehicle in front, and air bags, are all examples of automatic systems that rely on situational awareness. They can be found on modern vehicles regardless of whether the vehicle as a whole is capable of any degree of autonomous self-driving.

Regardless of the technological sophistication of the system, accurate and complete data, communicated and processed in a timely manner, is the foundation of accurate situational awareness.

Managing misuse

Reasonably foreseeable misuse is within the scope of ISO 21448. While at first blush this may seem counterintuitive, the key words here are *reasonably foreseeable*. Common examples could include behaviors typical of those who don't treat vehicles well, including overloading the vehicle, using it to push or pull in ways that it was never designed for, accidental bumps and scrapes, aggressive acceleration or braking, or attempting to drive it too fast for conditions. Car rental and fleet leasing companies could prove to be a good source of historical information regarding vehicle misuse, as renters and employees might be prone to being less careful with a vehicle that they personally do not own. This behavior, while disappointing, is foreseeable. So, it has to be accounted for.

In instances where vehicle decision making can lead to safety hazards, then the operation or assistance of a vehicle by a remote user, or via communication with a back office, is also within the scope of ISO 21448.

The impact of proximity

The proximity between the driver and the vehicle neither guarantees proper use, nor infers a probability of misuse. But proximity can be a contributing factor to a reduction or loss in situational awareness. In real life, a driver backing out of a tricky tight spot can quickly readjust their mirrors, turn their head in a different direction, or even roll down their tinted windows at night to listen for clues and gain better visibility in the darkness. In these actions, there is no delay. But a remote user only knows what the system is *telling* them, limited in scope to the accuracy and capabilities of the remote sensing equipment, and the communications connection between the two.

Regardless, a vehicle has to be safe no matter if it is being driven by a human on board, a human connected remotely, or via some degree of automation.

The References

Safety-relevant topics addressed by other standards

Some of the topics that are out of scope for ISO 21448, are addressed in other standards. These can be digested cover-to-cover, and they might also be cited as normative references (see next section).

Derived from ISO 21448:2022, Table 1 — Overview of safety relevant topics addressed by different standards[1]

Source of hazard	Cause of hazardous events	Within scope of
System	E/E system faults	ISO 26262 series
	Functional insufficiencies	ISO 21448
	Incorrect and inadequate Human-Machine Interface (HMI) design (inappropriate user situational awareness, e.g. user confusion, user overload, user inattentiveness)	ISO 21448 European Statement of Principles on human-machine interface [3]
	Functional insufficiencies of artificial intelligence-based algorithms	ISO 21448
	System technologies EXAMPLE—Eye damage from the beam of a lidar.	Specific standards EXAMPLE—IEC 60825
External factor	Reasonably foreseeable misuse by the user or by other road participants	ISO 21448 ISO 26262 series
	Attack exploiting vehicle security vulnerabilities	ISO/SAE 21434
	Impact from active infrastructure and/or vehicle to vehicle communication, and external systems	ISO 21448 ISO 20077; ISO 26262 series, IEC 61508 series
	Impact from vehicle surroundings (e.g. other users, passive infrastructure, weather, electromagnetic interference)	ISO 21448 The ISO 26262 series ISO 7637-2, ISO 7537-3 ISO 11452-2, ISO 11452-4, ISO 10605 and other relevant standards

Normative references in the standard

In the world of technical documentation, a user can quickly become overwhelmed by a dizzying array of standard numbers and citations. Accuracy and completeness are paramount, so some of this is unavoidable. But to make the standards a bit easier to read and consume, normative references are used to help strike a practical balance.

Simply put, normative references are any other documents which are referenced within a standard. These typically refer to established standards that were previously published and vetted by recognized groups, or publications that work in parallel to that standard.

In ISO 21448, normative references are referred to in the text in such a way that some or all of their content also constitutes requirements for ISO 26262. This is convenient for the user of the document. The authors of the standard have already ensured continuity between the documents, making it much easier for the reader to digest the presented content verbatim without having to constantly hop back and forth between two different standards documents that were deliberately written to work together in the first place.

Examples of references

For dated references in ISO 21448, only the edition cited applies. For undated references, the latest edition of the referenced document applies, including any amendments.

- By far, the most commonly-used normative reference in ISO 21448, is ISO 26262, *Road vehicles — Functional safety*
- For instances where a specific subsection of a standard is being referenced, it is cited as follows:

  ISO 26262-1, *Road vehicles — Functional safety — Part 1: Vocabulary*
- An example of a dated reference would be:

  ISO/PAS 21448:2019, *Road vehicles — Safety of the intended functionality*

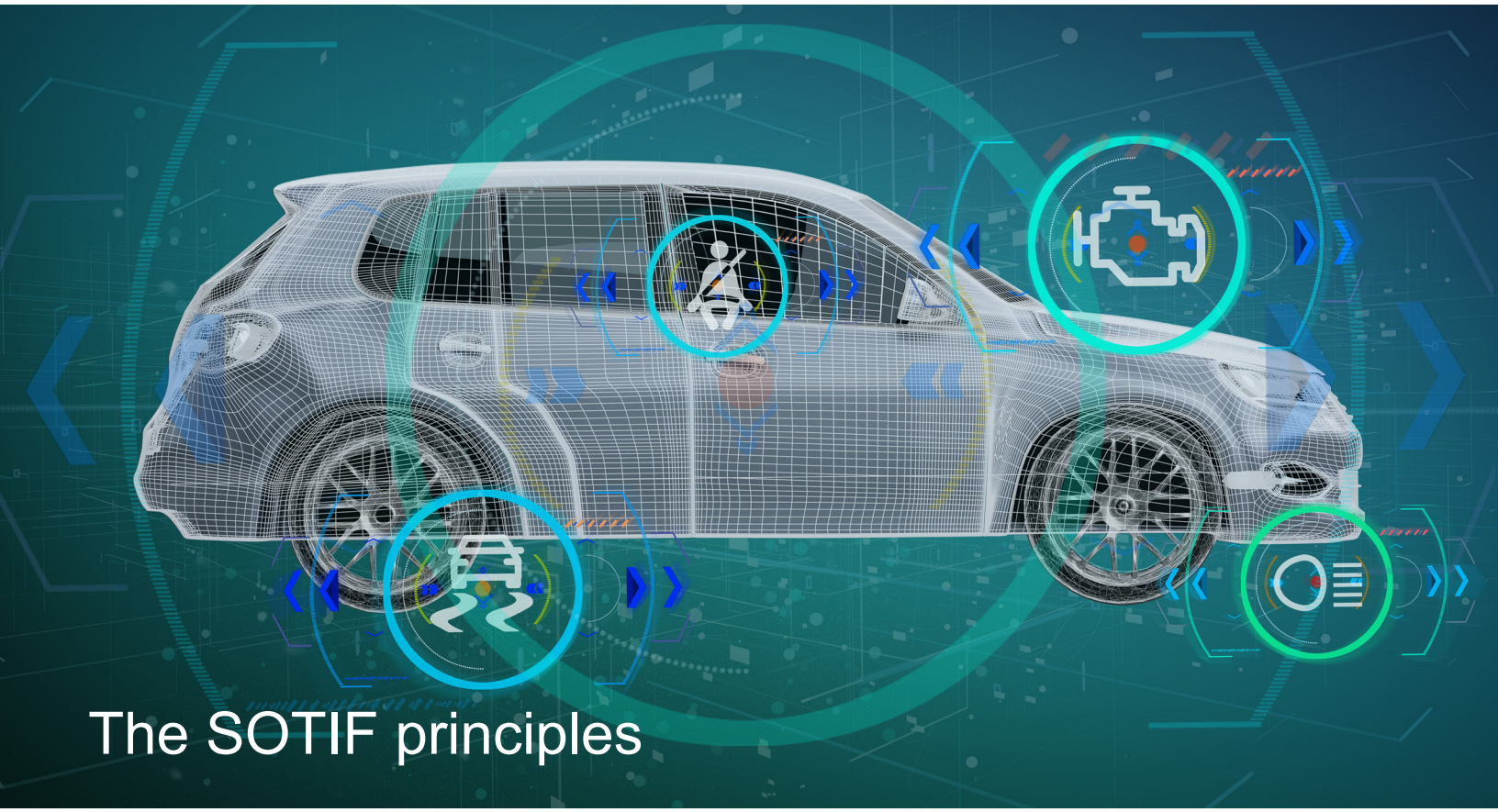
Terms, definitions, and external sources

Whether you are exploring topics that are new to you, or verifying prior knowledge, it is helpful to have terms and definitions at your fingertips that provide quick learning and confirmation. These might be found within the standard, in other standards, or in other online sources or databases.

Within ISO 21448:2022, *Clause 3. Terms and definitions*, there are 34 key definitions that include detailed examples and illustrations. Before you delve further into the standard, it is strongly recommended that you first become familiar with these terms and concepts, which provide a solid foundation for understanding the basics of the SOTIF, as well as further learning.

In ISO 21448, three external sources are also referenced. One is another standard, and the other two are terminology databases used in standardization and maintained by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC):

- ISO 26262-1, *Road vehicles — Functional safety — Part 1: Vocabulary*
- ISO Online browsing platform: available at https://www.iso.org/obp
- IEC Electropedia: available at https://www.electropedia.org/



The SOTIF principles

At its most fundamental level, SOTIF is based on an accurate assessment of the level of risk in a given scenario or event. This is attained by adherence to SOTIF principles, by following a defined workflow of SOTIF activities, and by the proper management of SOTIF activities and their supporting processes. These activities are applicable at the vehicle, system, and component levels.

The SOTIF-related hazardous event model

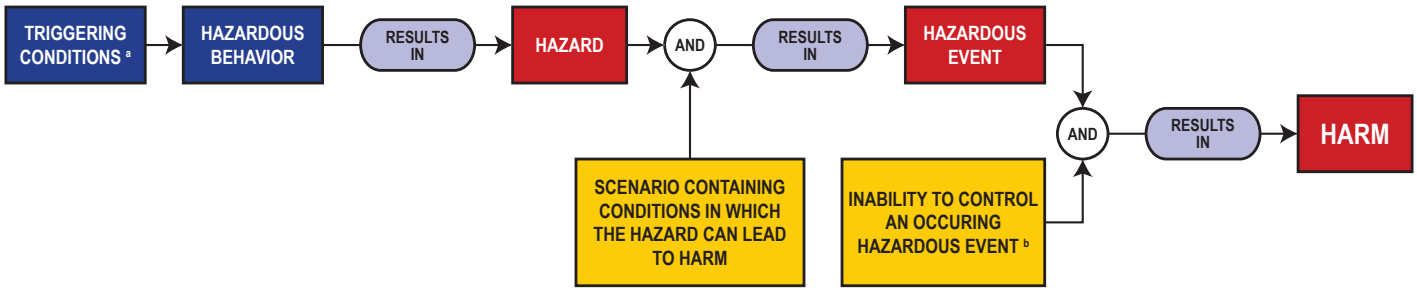
The main purpose of ISO 21448 is to describe the activities and rationale that are used to ensure that the level of risk associated with all of the SOTIF-related hazardous events that have been identified, is sufficiently low. These assessments take place in specific steps.

Vehicles are designed to fulfill a purpose. Therefore, the function, system specification, and design of a vehicle are shaped by relevant use cases. These use cases drive the definition of one or more scenarios. Scenarios are the fundamental building blocks of the SOTIF process, the information containers in which hazards, triggering conditions, and harmful results are defined, classified, and clarified.

A scenario could contain benign information that is factual but does not lead to harm. Or it could contain triggering conditions that lead to harm; In order to avoid this harm, proper situational awareness is necessary.

In their order of occurrence, these assessments and events take place as follows:

Derived from ISO 21448:2022, Clause 4.2.1., Figure 4 — Visualisation of a SOTIF-related hazardous event model[1]



Key: a. Triggering conditions include direct misuse that is reasonably foreseeable.

b. The inability to control the hazardous event can also be the result of a reasonably foreseeable indirect misuse, for example, if the driver does not supervise the system as he/she is supposed to.

EXAMPLE 1 — A functionality is designed into the vehicle, which is intended only for use on the highway at highway speeds and traffic densities. However, it is discovered that this functionality gets overwhelmed and has difficulty recognizing and properly interpreting the varying speeds, motions, and visual imprints of the wide variety of vulnerable road users typically found in dense urban conditions filled with roads, sidewalks, and crosswalks. This environment can include non-motorized road users like pedestrians and cyclists, persons with disabilities or reduced mobility and orientation, as well as motorcyclists.

EXAMPLE 2 — The driver interprets the operating mode of the system incorrectly, assuming that the system is active when it is actually deactivated. In a situation like this, potential insufficiencies in the system's human-machine interface could be considered hazardous behavior if the system fails to prevent confusion. Also, if the driver behavior can be monitored but there is an absence of an appropriate system reaction when the driver does something wrong or the driver doesn't react when they should, that can also be considered a hazardous behavior of the system.

Proper situational awareness

Situational awareness is the perception of elements and events in the immediate environment with deference to space and time, the accurate comprehension of their meaning, and the accurate forecast of their future meaning and impacts. It is both adaptive, and externally directed. **Situational awareness is a critical yet elusive foundation for decision making.**

Achieving proper situational awareness relies on a number of factors:

- **A sufficiently accurate and comprehensive perception** of the relevant environmental conditions, including:
    - An accurate understanding of the scene.
    - An accurate and complete forecast model for the state of each road actor; in other words, the model ensures that the things in the environment are accurately identified and their likely behavior is accounted for.
    - Localization, which is the adaptation of the vehicle to the local language and culture, and its traffic laws and control devices:
        - EXAMPLE 1: A vehicle that, when driven from the United States to Canada, automatically adapts from U.S. miles-per-hour and traffic control devices, to Canadian kilometers-per-hour and Canadian traffic control devices.
        - EXAMPLE 2: A vehicle that adapts as it travels from a state where traffic signals tend to be oriented vertically, to one where they are oriented horizontally.
    - Egomotion, defined as the 3D motion of a camera or other sensing devices within an environment, enabling a vehicle to estimate its motion as it relates to a rigid scene. An example of an egomotion estimation would be estimating a vehicle's moving position relative to the lines on a road or the street signs being observed from the car itself.
    - Communication with other vehicles or the environment around the vehicle.
- **The appropriate actions and reactions** while the vehicle is driving, including obeying traffic rules and the rules dictated by traffic signage.

Changes to situational awareness

Situational variations

Vehicles and the environments they operate in, are in a state of continuous flux. The characteristics of the vehicles themselves can change due to wear, maintenance (or the lack thereof), hardware and software upgrades, and even changes to the fuels and lubricants they use. The loads they carry, and where that weight is placed in the vehicle, can also cause changes in the vehicle's operating characteristics, albeit temporarily. And of course, the environments they operate in are constantly changing due to weather and physical differences from location to location.

These are examples of some of the elements that may vary over the life of the vehicle:

- The environment in which the vehicle is driving may vary, including:
  - temporary lane adjustments and detours due to road construction and repair;
  - new types of traffic signs;
  - new traffic laws;
  - new roads, intersections, entrances, exits, pedestrian crosswalks, railroad crossings, and additional lanes;
  - new driving scenarios;
  - changes in existing driving scenarios; and
  - weather.
- Appropriate reactions may vary, including:
  - New driving actions required by any of the changes in the environment listed above;
  - changes in driving laws; and
  - changes to vehicle handling due to weight distribution.

Managing changes to situational awareness

At first blush, monitoring all these variations may seem overwhelming. And it is true that there are a lot of considerations that need to be covered. However, there is a vetted process for doing so. These changes are monitored by adherence to the procedures and activities addressed in ISO 21448:2022, Clause 13, *Operation phase activities*[1]. This process is completed before any formal systems updates are released.

Changes that impact the driving policy

The driving policy is the implementation of the vehicle-level SOTIF strategy (VLSS) at the decision-making level. For example, the requirements for the vehicle's transition from a normal state to a degraded state due to departures from the operational design domain or the erosion of expected performance, are within the scope of the VLSS. EXAMPLE: When the system places an engine into derate mode and limits the vehicle speed to a 30 mph "limp home" maximum due to a catalytic converter failure.

Situational variations can have a significant impact on the driving policy. In ISO 21448, the design and specification of the driving policy is addressed in Annex D (informative), *Guidance on specific aspects of SOTIF*[1], and Annex D.1 *Guidance for driving policy specification*[1].

Annex D.1 provides detailed guidance as to how a driving policy can be designed, and it provides examples of an implementation.

As a result of performing the processes defined in Annex D.1, changes to situational awareness are identified and addressed. And as time goes on, the effectiveness of the driving policy can be evaluated by comparing it to statistics gathered about the usage of the system, to help determine whether further refinements and another round of system updates are called for.

The SOTIF scenarios

Note: In previous versions of ISO 21448, the terms “Safe” and “Unsafe” were used, where “Not Hazardous” and “Hazardous” are used in the latest version, published in mid-2022. We reference the 2022 version of the standard in this eBook, but as of this writing, the older terms are still in common use in the industry, and in previously published works. For the intent and purpose of this conversation, their relative meanings are basically the same.

Vehicles are designed to fulfill the requirements of one or more use cases. In order to select the proper next steps to address the considerations in a scenario, the first step is to determine what type of scenario it is. For example, a scenario that causes hazardous behavior is referred to as a hazardous scenario.

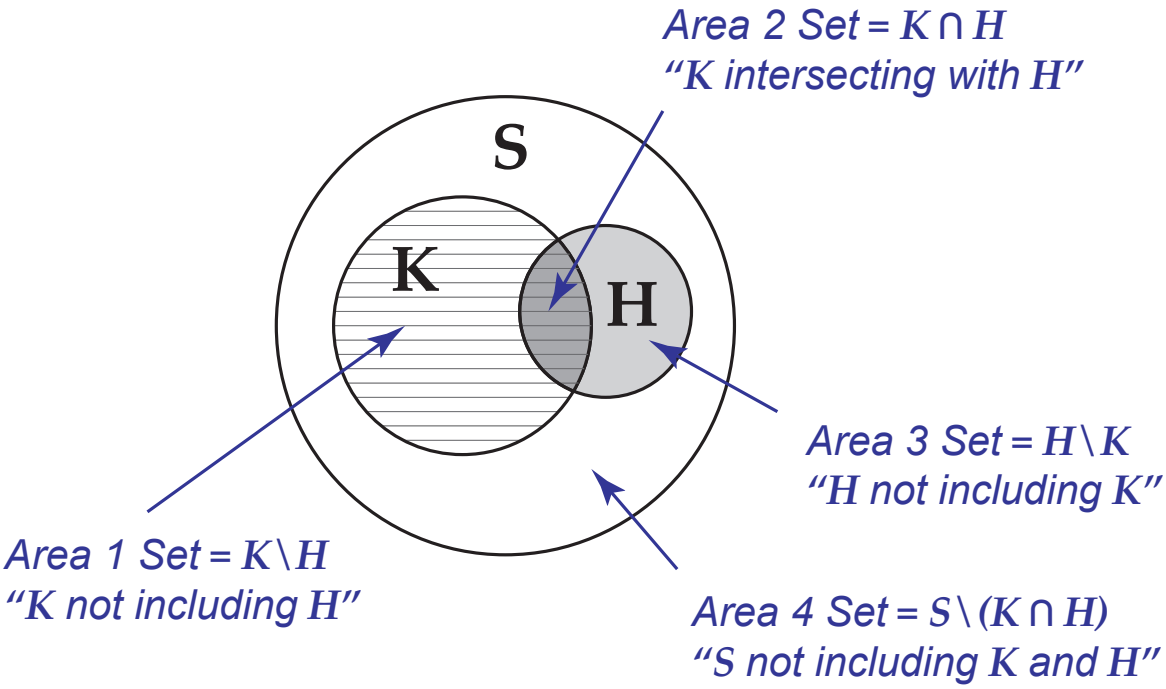
SOTIF scenarios which are part of relevant use cases are classified into one of four areas and keyed to specific numbers:

1. Known and Not Hazardous
2. Known and Hazardous
3. Unknown and Hazardous
4. Unknown and Not Hazardous

In ISO 21448:2022, these scenarios are represented visually in two ways.

In Figure 5 (next page), the visualization depicts a typical example of the approximate overlap and relative quantity of scenario areas according to how they are categorized, and how they overlap with each other:

Derived from ISO 21448:2022, Clause 4.2.2., Figure 5 — Visualisation of scenario categories[1]

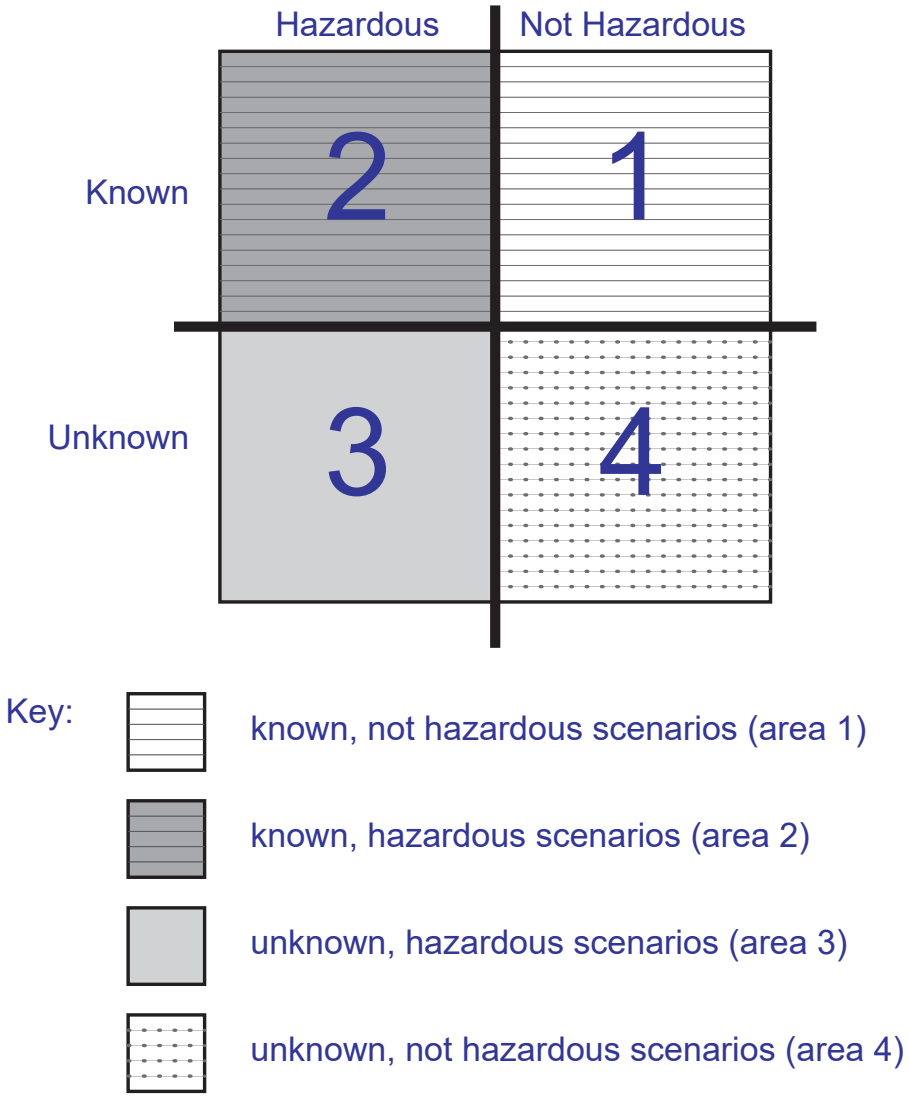


Key:

- represents the set of known scenarios **K**
- represents the set of hazardous scenarios **H**
- represents the set of all possible scenarios **S**

This visualization of the logical relation between the four scenario areas, rendered in a Venn diagram style, is used in conjunction with Figure 6 (next page), which is a four-box style of graphic that keys the type of scenario to a specific number:

Derived from ISO 21448:2022, Clause 4.2.2., Figure 6 — Alternative visualisation of scenario categories[1]



In both Figure 5 and Figure 6, the size of the areas represents the approximate quantity of scenarios relative to the total number of scenarios, not the degree of risk. The severity of resulting harm, and likelihood of occurrence, both contribute to the risk of the intended functionality, but are not represented visually in these graphics.

Evaluations and probabilities

Figure 6 is a conceptual abstraction that represents the goal of SOTIF activities, namely, the avoidance of unreasonable risk. This is achieved through the following activities:

- Performing an evaluation of the acceptance of risk in the known and hazardous scenarios found in area 2, based on the intended functionality.
- Using functional modification, reducing the probability that the known hazardous scenarios in area 2 will cause hazardous behavior.
- Using an adequate verification and validation strategy, reducing the probability that the unknown hazardous scenarios in area 3 will cause potentially hazardous behavior.

The ultimate goal of SOTIF activities is to thoroughly evaluate the potential hazards in areas 2 and 3, and to demonstrate that the residual risk caused by these scenarios is at or below the acceptance criteria.

Area 1 — These risks are known and not hazardous. There is no significant mystery about them, and no corrective action is required. They are acknowledged and monitored as a part of the typical SOTIF review processes, just to make sure that nothing about them has changed to cause them to be moved to a different category.

Area 2 — These risks are known and hazardous. These risks are explicitly evaluated. These are the risks that carry the least mystery. The risks are known, so the team knows right where to start. And the risks are known to be hazardous, so there is no question that they need to be addressed.

Area 3 — These risks are unknown and hazardous. These are the risks that can bite unexpectedly. The risks associated with these unknown scenarios are assessed using statistics-based testing. And of course, once they become known, they are recategorized to Area 2.

Area 4 — These risks are unknown and not hazardous. Similar in safety impact to the risks in Area 1, no corrective action is required even though these risks are unknown. And once they become known, they are recategorized to Area 1.

Unknown hazards versus known hazards

A given use case can contain both known and unknown hazards in its scenarios.

It is important to thoroughly explore all the potential scenarios of each use case. This exploration can lead to the discovery of previously unknown scenarios; as they are resolved, risk can be further reduced and safety improved.

Unknown hazards

A scenario is classified as unknown when:

- The behavior of the system is unknown, even if potentially triggering conditions have been identified. EXAMPLE: The effects of extremely cold weather creating a cold-soaked vehicle.
- There are unknown triggering conditions.
- The known parameters of different scenarios can combine, and form unknown potential triggering conditions. EXAMPLE: A combination of specific weather and traffic conditions in a particular geographic location, such as high wind and snowfall conditions on a steep twisting mountain pass in winter.

Known hazards

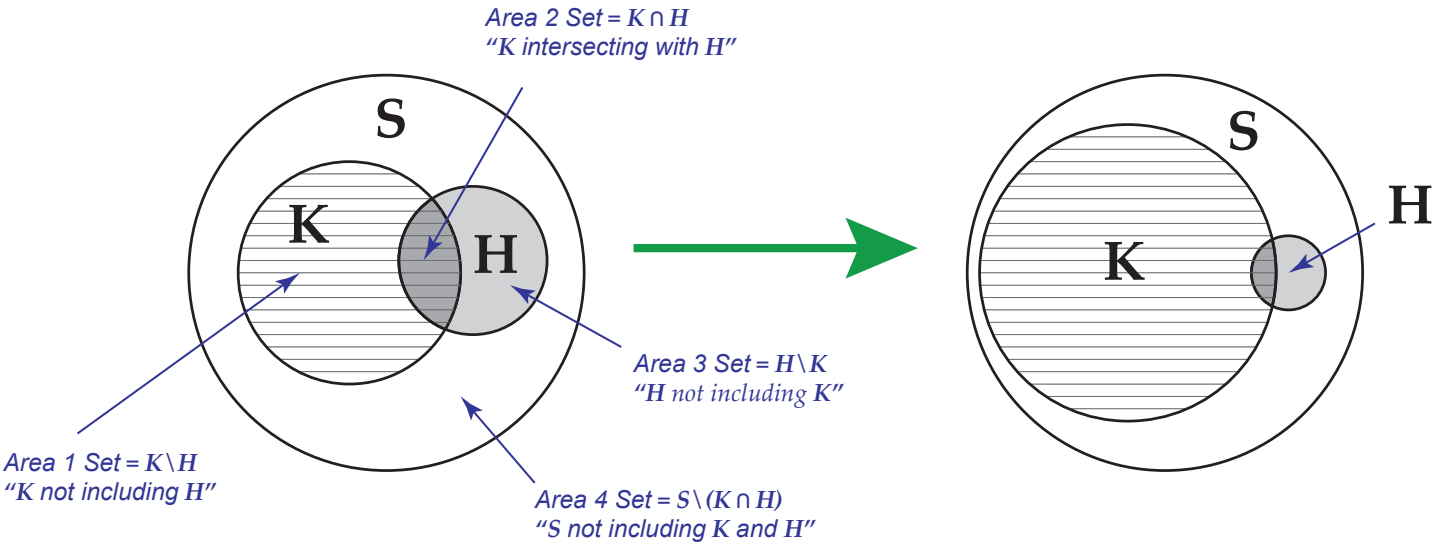
Hazardous scenarios fall within the scope of SOTIF regardless of wether they are known or unknown. (After all, a hazard exists whether we know about it or not.) But once a hazard becomes know, it must be dealt with.

Scenarios in Area 4 that are unknown but not hazardous, do not carry a risk of harm. But once a scenario in Area 4 becomes known, it is recategorized to Area 1 as now being known but retains its classification as being not hazardous.

Evolution and progress

How can you tell that progress is being made? The expectation is that the risks resulting from Areas 2 and 3 will be reduced, as these risks are identified and their challenges solved. As more scenarios progress into Area 1, confidence will increase that the SOTIF is being achieved. This progress can be expressed using both the Venn diagram and four-box style of models:

Derived from ISO 21448:2022, Clause 4.2.2., Figure 7 — Evolution of the scenario categories resulting from the ISO 21448 activities[1]



Example of an initial starting point of development

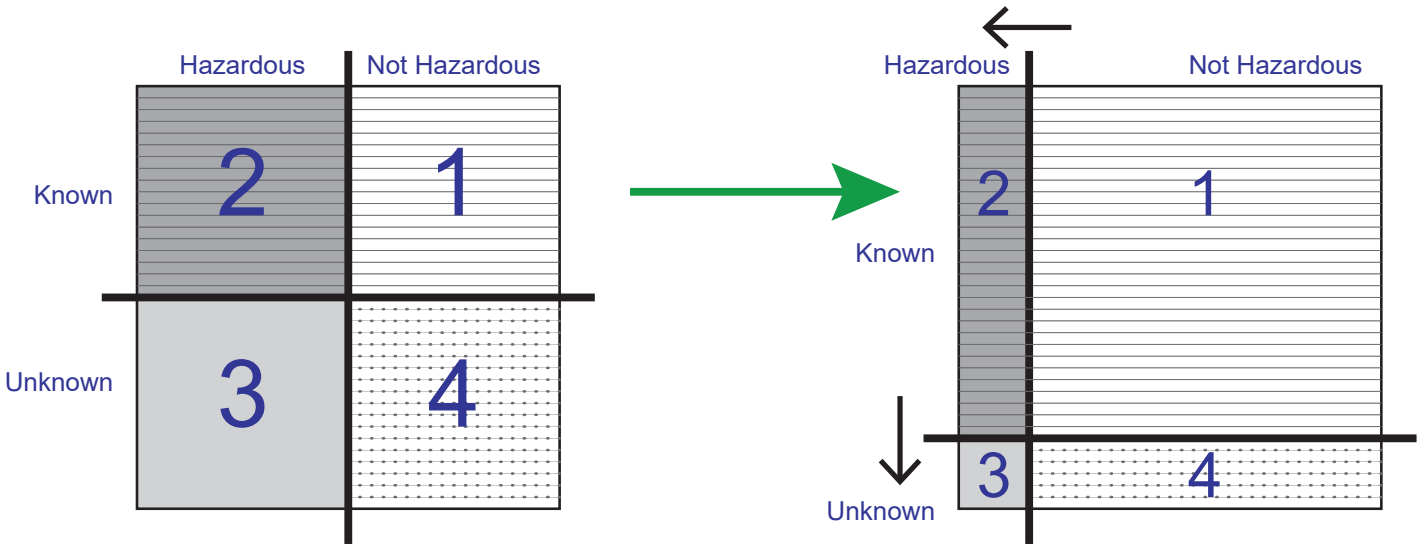
Goal for the SOTIF release

represents the set of known scenarios **K**

represents the set of hazardous scenarios **H**

represents the set of all possible scenarios **S**

Derived from ISO 21448:2022, Clause 4.2.2., Figure 8 — Alternative evolution of the scenario categories resulting from the ISO 21448 activities[1]



Example of an initial starting point of development

Goal for the SOTIF release

Key:

- known, not hazardous scenarios (area 1)
- known, hazardous scenarios (area 2)
- unknown, hazardous scenarios (area 3)
- unknown, not hazardous scenarios (area 4)

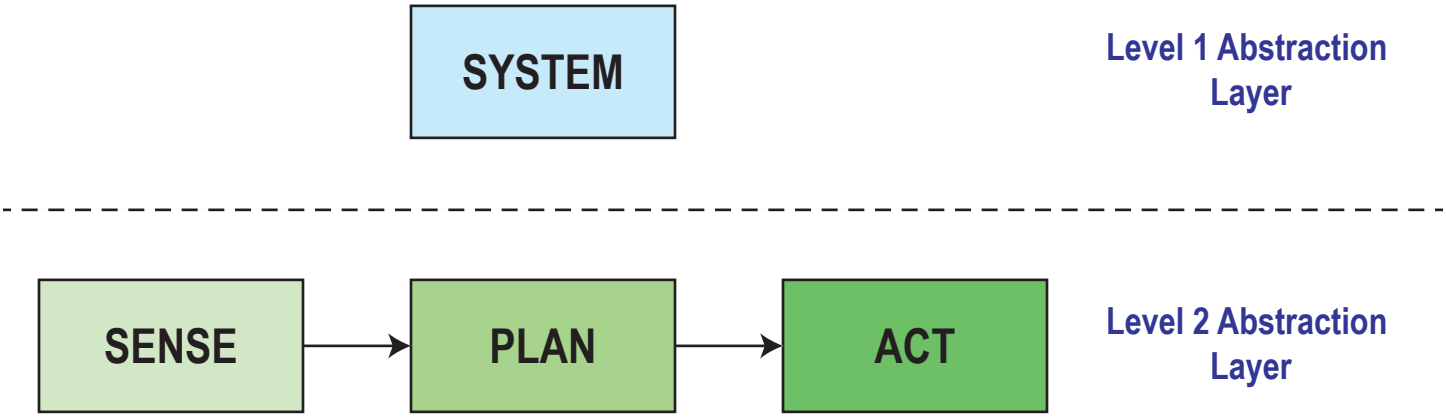
The Sense-Plan-Act model

The possible causes of hazardous behavior that are considered within the SOTIF standard, are closely linked to the system's ability to:

- create a model of the environment that is sufficiently accurate and complete;
- make the correct decisions;
- determine the correct control actions based upon that environmental model; and then
- properly execute those control actions.

The key system elements and their interactions are represented by the "Sense-Plan-Act" model, illustrated below:

Derived from ISO 21448:2022, Clause 4.2.3., Figure 9 — Visualisation of the Sense-Plan-Act model[1]



How does the Sense-Plan-Act model work?

The model consists of four key system elements (represented by the rectangular boxes), one on the higher first level of abstraction, and the remaining three on the lower and more detailed second level of abstraction.

Decision algorithms are included in all of the key system elements in the Sense-Plan-Act model. These can include algorithms for sensor data, classification, fusion, analysis of the situation, and descisions about what actions to take.

Following the model's flow from left to right:

1. The key system element "Sense" performs the **perception activities**, where an environmental model is created based on the information that is received when the system senses the vehicle's internal environment, external environment, vehicle state, and system state. This model includes appropriate localization.

2. The key system element "Plan" is equipped with **goals and policies** and applies them to the environmental model generated by the "Sense" element. From this activity, the control actions are derived.

3. The key system element "Act" then **executes the control actions** generated by the "Plan" element.

A key part of the system architecture

The selection of a capable and comprehensive system architecture is an important consideration when striving to achieve an efficient SOTIF process. The Sense-Plan-Act model provides a solid framework upon which the prescribed activities can take place, both at the early stages of development, and throughout the entire functional development lifecycle.

Selecting a competent and capable system architecture is critical for ensuring the SOTIF. Therefore, the activities surrounding the definition of the system architecture should be started early in the system development process.

How ISO 21448 is used to achieve the SOTIF

The detailed order of SOTIF activities

The SOTIF activities are performed in a specific order:

1. **The specification and design are defined.** They include functional insufficiencies that are already recognized prior to the downstream SOTIF activities and cycles. Each iteration that starts from the specification and design, relies on the specification and design first being brought up to date. SOTIF activities should always be started from the latest known-good data.

2. **A hazard identification and risk evaluation are performed on the potentially hazardous behaviors of the intended functionality.** The identified hazardous events are evaluated based on their risk. Then, their risk acceptance criteria are defined accordingly. Compare and contrast the following:

- If the hazardous events **do not** lead to unreasonable risk, then no additional design measures are applied, as there is no safety-based need for them.
- For the intended functionality, the causes of hazardous behavior are not considered, only their **consequences** for safety. The purpose is to evaluate hazardous events that could result from hazardous behaviour, and to then define the acceptance criteria that must be met.

3. **The possible root causes for the hazardous behaviors are evaluated.** It is then determined whether the risk posed by potential functional insufficiencies and triggering conditions is logical and practical.
4. **The functionality is modified to improve the SOTIF,** if it is deemed neccesary.
5. **A verification and validation (V&V) strategy is applied.** The V&V provides evidence that:

   a. the vehicle-level risk is below an acceptable level,

   b. the elements meet their functional requirements, and

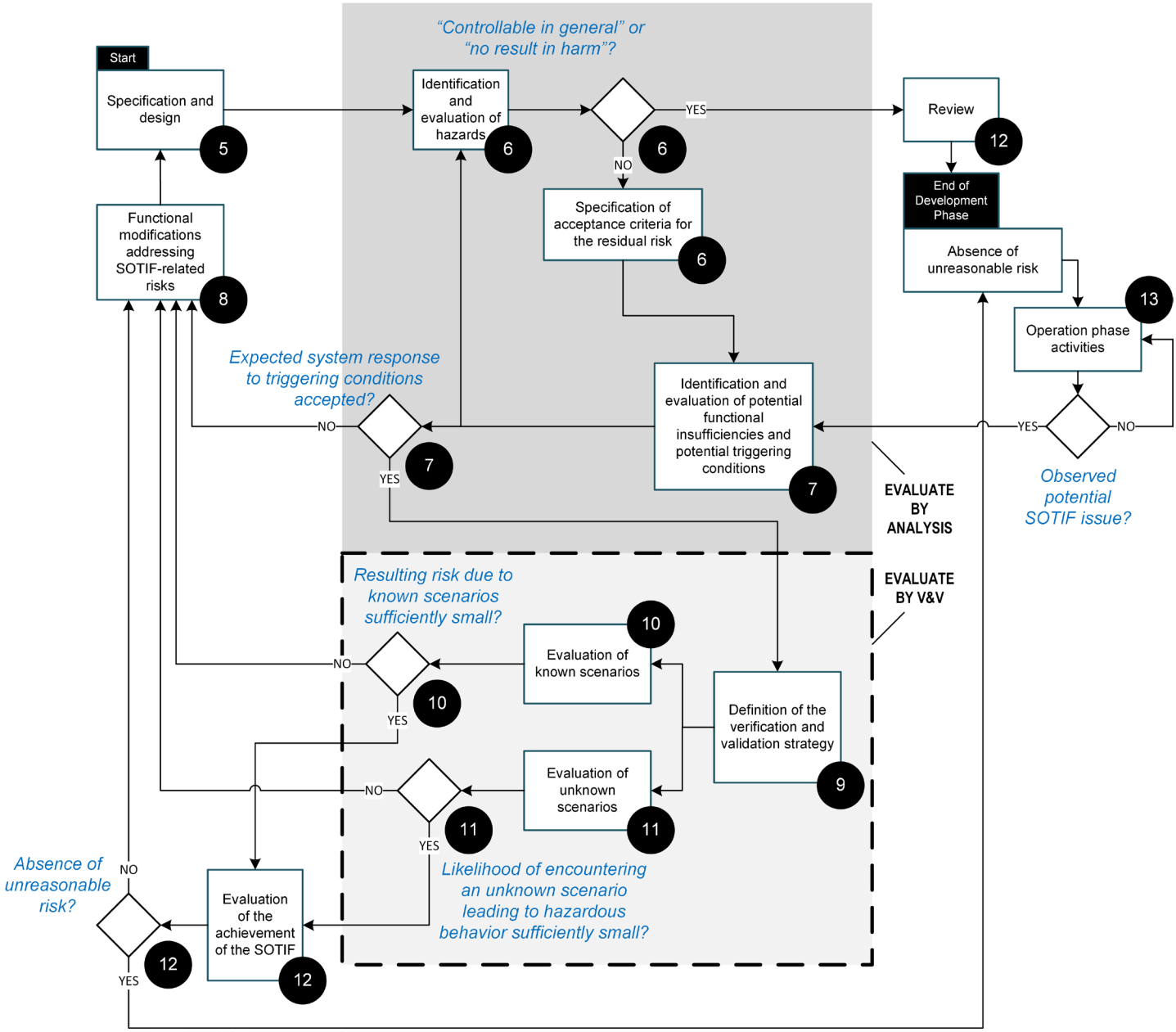
   c. the domain of the operational design is satisfactory.

6. **The required evidence is collected.** Corresponding V&V test cases are derived, and it is confirmed that the test case coverage over the operational design domain is sufficient.
7. **The results of the SOTIF activities are evaluated,** to determine whether the achievement of the SOTIF can be successfully demonstrated.
8. **The operation phase includes a process to evaluate and resolve possible SOTIF issues that emerge during the course of field operations.** It is unlikely that all issues will be identified beforehand or at the first attempt. The processes in this phase capture new knowledge and lessons learned, and turns them into actionable improvements.

The flow of SOTIF activities

ISO 21448:2022, Clause 4.3.1., Figure 10, describes the flow of the activities that are required to ensure the safety of the intended functionality. Note that there are different groups of activities for evaluation by analysis, versus evaluation by V&V. The circled numbers denote the corresponding clauses within the ISO 21448 document.

Derived from ISO 21448:2022, Clause 4.3.1., Figure 10 — Dependencies between the ISO 21448 activities[1]



The Annexes in ISO 21448

The annexes in ISO 21448 provide a wealth of rich detailed information and explanations, examples, and reference information for key aspects of the standard, often visually illustrated with helpful graphics, tables and flowcharts. **The annexes are summarized as follows, and their review is strongly recommended:**

Annex A provides informative general guidance on SOTIF:

- Structuring the SOTIF argument using goal structuring notation (GSN), a method widely used in the safety community.
- Explanations regarding the interaction between functional safety according to the ISO 26262 series and ISO 21448:2022, including possible interactions of product development activities between the two documents.

Annex B provides guidance on scenario and system analysis, including:

- The method for deriving SOTIF misuse scenarios.
- The construction of scenario factors for the SOTIF safety analysis method.
- Examples of the adaptation of safety analyses to identify and evaluate potential triggering conditions and functional insufficiencies.
- Applying System-Theoretic Process Analysis (STPA) in the context of SOTIF for ADAS and automated vehicles.

Annex C provides guidance on SOTIF verification and validation, including:

- The purpose of the verification and validation strategy.
- Derivation of validation targets.
- Validation of SOTIF applicable systems.

- Perception system verification and validation.
- Guidance on scenario parameterization and sampling.
- Considerations for reducing validation testing.

Annex D provides guidance on specific aspects of SOTIF, including:

- Guidance for driving policy specification.
- Implications for machine learning.
- SOTIF considerations for maps.
- SOTIF considerations for vehicle-to-everything (V2X).

The management of the SOTIF activities and their supporting processes

Rigorous engineering and quality management processes are paramount for developing a safe product. Through these processes, the specification and design of the system are defined, and the potentially hazardous behaviors of the intended functionality are evaluated. There are many types of processes, and they happen in a carefully defined sequence.

Parallel activities between the standards

Processes and activities specified in ISO 21448 and the ISO 26262 series can be carried out in parallel. Because implemented measures in general can have an impact on SOTIF as well as functional safety, they are evaluated by both of the disciplines.

Special attention should be paid to the cascading of requirements, and traceability.

SOTIF development activities for distributed product development

A **distributed product development** is a condition where the product is developed by two or more different companies or organizations that have entered into a formal business relationship. In the case of a distributed product development, special care is required to coordinate and streamline efforts to reduce variation and waste, and to help protect the interests of the respective parties.

A **development interface agreement (DIA)** must be defined between all involved parties. The goal of the DIA is to confirm, in the early stages of a project, all of the responsibilities of the SOTIF activities. Among other benefits, this important step helps all the respective parties to determine whether or not they have the right people in the right place at the right time. The DIA also helps to ensure that adequate technical information will be exchanged between the development parties.

Completion of the DIA is a fundamental but critical early foundational step that must be completed before all the other subsequent downstream work can take place.

SOTIF-related elements and their relationship to context

To achieve the SOTIF, it is critical that the interfaces between the hardware systems and software systems be described accurately and completely, to include capturing their proper context with each other.

Thus, the *boundaries* of each system must also be carefully evaluated. Because the environmental factors have an essential impact on the issues surrounding development of the SOTIF, the systems and their elements may have different concerns in this regard, depending on their position within the hierarchical layers.

The development of these systems and elements can be categorized as being one of the following three types:

- **In-context development:** The complete system is developed using all of the SOTIF activities detailed in the V model.
- **SOTIF-related elements that are out of context:** Assumptions can be made regarding the use of these elements within the whole system and their impact upon the intended functionality. These assumptions must be documented. They are then used as inputs for the subsequent development of the SOTIF-related elements. In turn, the validity of the assumptions is established by SOTIF activities within the context of all the vehicle-level functionalities.
- **Non-specific SOTIF-related development:** This can be thought of as the SOTIF equivalent of “it depends”. The functionality of non-specific elements can contribute in so many different ways to the intended functionality, that it is essentially not realistically feasible to estimate their priority without first knowing the context in which these elements will be used.

The specification and design of the SOTIF

The specification and design of the SOTIF contains all of the information required to conduct the SOTIF-related activities. After each iteration of the SOTIF-related activities, the specification and design are updated as required.

The specification and design can include a wide variety of aspects and considerations that can be approached in different ways. Some are relevant only for a specific implementation or level of automation. Likewise, some aspects are relevant at the level of the entire vehicle, while others are relevant only at the level of a specific element.

ISO 21448:2022, Clause 5.2 *Specification of the functionality and considerations for the design*[1], provides a detailed list of possible aspects for consideration, including those that impact:

- the intended functionality;
- the functionalities of supporting subsystems and components;
- the design of the systems and elements implementing the intended functionality;
- the performance targets of the installed sensors, controllers, actuators, and other componenets and inputs;
- the performance targets of automated driving systems, including their detection capabilities and responses to critical objects and events;
- the intended functionalities' dependencies on, and interactions or interfaces with:
    - the driver of the vehicle;
    - the interface being utilized by the driver;
    - how the interface is being used to mitigate misuses that are known and reasonably foreseeable;
    - remote operators and back office operators;

  - passengers, pedestrians, cyclists and other users who share the road with the vehicle;
  - road infrastructure and roadside objects used for the safety and control of traffic;
  - road infrastructure and roadside objects used to assist the driver;
  - the exchange of data to and from the cloud and the vehicle;
  - inter-vehicle communications or other communication infrastructures;
  - in-service telematics involving diagnostics and parameter updates;
  - the remote flashing of software updates;
  - relevant environmental conditions, even as they constantly evolve;
  - the other functions of the vehicle that might interfere with the intended functionality, including the exchange of information, and any corresponding assumptions that may arise about its use;
- reasonably foreseeable direct and indirect misuse;
- potential performance insufficiencies, identified triggering conditions and any system or element countermeasures;
- the system and vehicle architectures that are used to carry out the intended functionality;
- the various warning and degradation concepts, strategies, and schemes;
- the procedures that are utilized for monitoring and data collection; and
- the mechanism, design and requirements that bolster the risk mitigation abilities during the operation of the vehicle.

Considerations for the design of the system and its architecture

The importance of an all-encompassing description of the system

Achieving understanding

The purpose of the specification and design is to provide teams with an adequate understanding of the system, all of its elements and functionality, and the targets that have been defined for its performance.

This comprehensive understanding is neccesary so that the subsequent activities in all of the phases can be properly completed. This understanding must include a throrugh and detailed list of all the known functional insufficiencies, any related triggering conditions and, where applicable, the appropriate countermeasures.

Continuous improvement

Some of these potential issues and considerations are known and documented before the SOTIF-related process even begins. Others come to light as a result of the SOTIF activities. The process brings them together at each iteration, so that each iteration is a snapshot, at one point in time, of the latest and greatest information.

As new functional insufficiencies and triggering conditions are identified during the SOTIF process, along with any relevant pre-existing content. Measures to improve the SOTIF are defined, and both the specification and design are updated each time that the development cycle is executed.

Each iteration of the SOTIF-related activity can result in engineering activities which can in turn drive updates to either the specification and/or the design, at any relevant level as required, so that each iteration encompasses all of the information previously discovered. This results in the system being designed in such a way that the latest countermeasures are thus woven into that iteration, to blunt and alleviate the effects that known insufficiencies might have on the overall system. This is how the prior knowledge and lessons learned become actionable, resuting in direct improvements to safety.

Methodically encompassing previous work in new updates

It is important to ensure that updates do not leapfrog previous work. The SOTIF work products are linked with the specification and design if they have any kind of impact on them, including any impacts to relevant pre-existing content. This linking can become quite complex, but it is absolutely vital that it be maintained accurately and completely. Thus at any given time, the entire team, regardless of who they are or which partner they work for, should all only be working on the next iteration..

Cooperation among the development parties

Broad, consistent, complete, and reliable communication and cooperation among all of the parties involved with the development of the vehicle system is necessary to discover potential insufficiencies and to develop appropriate countermeasures.

Shared information should flow upward and downward, and inward and outward, throughout the entire partnership hierarchy. The relevant sections of the design and specification are communicated to lower-level system and component developers as appropriate. Likewise, after each development cycle iteration, assumptions about use, foreseeable misuse, and the potential performance insufficiencies, are communicated upward from one tier to the next in the hierarchy, up to and including the OEM.

Traceability

While conducting the specification and design work, traceability and completeness can be demonstrated by linking to the SOTIF measures and work products. In turn, these can be further linked with:

- Relevant design documents.
- The work products from:
  - Evaluating the risk of hazardous behaviors.
  - Evaluatiing the system’s response to triggering conditions.

- The verification and validation results for hazardous scenarios that are known.
- The validation results for hazardous scenarios that are unknown.
- The arguments made for the release of the SOTIF, and any reasons for rejecting the release request.
- The processes used for monitoring the systems in the field, and any new hazardous scenarios discovered during this field monitoring.

Technical assumptions related to the evaluation of risk may not necessarily be associated with SOTIF measures, but they can still be traced to the specification and design. This helps future engineers to understand what the previous persons were thinking. Design tools that offer model-based design and supporting traceability between requirements, components, interfaces, analysis, test cases, and results, can support this process.

Performance insufficiencies and their countermeasures

The SOTIF design includes considerations on potential performance insufficiencies that can potentially result in vehicle-level hazardous behavior, including but not limited to:

- Insufficient classification, measurements, tracking, target selection, or kinematic estimation.
- False positive detections, commonly called “ghost” or “phantom” objects.
- False negative detections.
- Limitations at the driving policy level, such as the consideration of occluded areas.

Functional insufficiencies carry the highest relevance when the system operates within its specified operational design domain. But the way the system detects when it is leaving its specified operational design domain, and how it operates during those transitions, is also relevant to support an accurate and complete analysis.

The development of the system is based on assumptions that are made about the performance insufficiencies inherent in the design. To safeguard the SOTIF, measures are then put

in place to cope with these performance insufficiencies. The effect of integrating the design and measures into the specification and design, increases the overall robustness and decreases the residual risk. ISO 21448 also details the means of discovering the potential functional insufficiencies and their triggering conditions, including redundant, diverse, and complimentary elements.

The identification and evaluation of hazards

The hazards arising from the intended functionality must be identified in a systematic manner and defined at the vehicle level. Just as an engineering document is meant to be digested and applied cover-to-cover, a vehicle must likewise be dealt with as a whole.

Accuracy and completeness are paramount in this endeavor, and are defined by two key considerations:

- The risk that arises from the hazardous behavior of the intended functionality, and the corresponding scenarios in which the hazardous behavior can lead to harm, must be systematically identified and then evaluated.
- The parameters that define these circumstances, and the acceptance criteria for the residual risk, must also be specified.

Sources of hazard information

To achieve these objectives, various sources of information can be considered. The specification and design can be divided among or linked to several of the documents of the SOTIF-related systems, such as the:

- requirement specifications;
- functional specifications; and
- design specifications.

In addition, the mitigation measures can be integrated into existing functional safety design documentation such as the functional safety concept and/or technical safety concept documents. The available data related to derivations from the acceptance criteria can also be reviewed.

Identifying the hazards

The hazards resulting from insufficiencies in the function are determined at the vehicle level in a systematic manner, based primarily on knowledge about the function and any possible deviations that might result from the functional insufficiencies.

The common elements of a hazard analysis include:

- Occurrence
- *Exposure
- *Controllability
- *Severity
- Risk

**The parameters of severity, exposure, and controllability, can be used to adjust the validation effort.*

In their order of occurrence, the events that lead to identifying a hazard include:

- A scenario containing **triggering conditions,** which results in…
- **hazardous behavior,** which leads to…
- **the hazard itself** which, when combined with…
- a scenario containing **conditions in which the hazard can lead to harm,** leads to…
- **a hazardous event,** including those that are not controlled, which in turn leads to…
- **harm.**

An important difference exists between the occurrence of a triggering condition, and the exposure to a scenario in which the hazard could lead to harm. **In general, triggering conditions are not independent from scenarios.** Therefore, in order to utilize the exposure to a scenario within an argument for the reduction of risk, the evaluation must detail the statistical dependence between the *probability* of being in a scenario and the *probability* of encountering a triggering condition. Prioritization without the context of probability, is wasteful.

The evaluation of risk

It is quite possible to estimate the severity of harm and the controllability of hazardous events through the use of a risk evaluation, which is an analysis of the risk presented by the hazardous behavior in given scenario. This information helps to quantify the acceptance criteria inherent in a given SOTIF-related risk.

The severity and controllability of the hazardous event are considered to determine if the resulting risk is unreasonable in a given scenario. Exposure itself is not a determining parameter. As the risks are evaluated in the scenarios, their selection already implies that they could impact the SOTIF. Otherwise, they would not be considered for analysis in the first place.

The severity and controllability evaluation

The functional specification is taken into account during the evaluation of the severity and controllability of the risk. Unreasonable risk is determined to be absent if the controllability is asessed as being as “controllable in general” or the severity is asessed as “no resulting harm”. **In all other cases, a hazardous event is considered to be SOTIF-related.**

The corresponding hazardous behavior is then described using specific and measurable parameters such as deviations in speed, or the minimum distances between the vehicle and other objects. These measurements provide information that is actionable.

The evaluation of delayed reactions or the lack of reactions by people

When an attempt is made to control the hazard, the controllability evaluation takes into account “no reaction” or “delayed reaction” by persons involved with the scenario, whether they are the driver, or others. EXAMPLE: A delayed reaction resulting from reasonably foreseeable indirect misuse. Measures occuring externally to the vehicle, such as persons in proximity to the vehicle from the road or roadside, or persons who have direct or indirect remote impact on the operation of the vehicle, can also be considered by this evaluation.

The evaluation of controllability can be impacted by a delayed or inappropriate reaction by the driver, and this topic is a part of the SOTIF-related analysis. This includes the time that the driver needs to recover and gain situational awareness, and whether or not they can success-fuly recover at all.

Specification of the acceptance criteria for residual risks

Acceptance criteria defines whether or not a risk is hazardous. But first, the criteria itself needs to be evaluated to ensure that it is appropriate. ISO 21448 details the considerations for evaluating the appropriateness of acceptance criteria. This list includes:

- applicable regulations, both in industry and from the government;
- whether the risk is unreasonable for persons who might find themselves exposed to the risk;
- whether the function in question is new or has already been established in the market;
- the acceptance criteria of those functions that have already been established; and
- the performance of a driver who far exceeds the norm in a commendable way.

The available traffic data for the target market, and pre-existing information from similar functions already operating in the real world, can be taken into consideration. And appropriate quantitative acceptance criteria can be included, as long as a valid rationale can be provided.

Identifying and evaluating potential functional insufficiencies and potential triggering conditions

Potential insufficiencies and triggering conditions (including those originating from reasonably forseeable direct misuse) must be identified, and it must be determined as to which of those lead to hazardous behavior. In turn, the resulting response of the system must be evaluated to ensure that it is acceptable for supporting the SOTIF.

The analysis process is not just focused inward. When potential functional insufficiencies and triggering conditions are systematically analyzed, field experience and knowledge gained from similar projects or experts can also be taken into consideration. This is one of the many reasons why transparent communication among the business partners is so important, as relevant knowledge or experience might be found in any segment of the team.

Combining methods

Various combinations of methods can be utilized to identify and assess the potential insufficiencies that might be found in the specification, performance insufficiencies, triggering conditions, or output insufficiencies. The need for combining methods might be driven in large part by the hardware itself. For example, the system might employ two different sensors to control a system, such as a radar and a camera, that are fused together by a domain controller. In this instance, it is allowable for unique performance targets to be assigned to each contributing element.

Addressing multiple triggering conditions

There can be multiple triggering conditions that need to be accounted for, as well as foreseeable misuse or known environmental conditions. This is an example of where traceability can play such an important role. Traceability threads together the hazardous behaviors, triggering conditions, potential performance insufficiencies, and any insufficiencies in the specification itself at either the vehicle or the element level. (In order to maintain a more structured and useable presentation, planning algorithms, sensors and actuators are handled separately.)

Analysis of reasonably foreseeable misuse, either direct or indirect

Reasonably foreseeable direct and indirect misuse of the intended functionality can serve as a wild card and contribute to an unreasonable level of risk. There can be many causes, including:

- the user lacking an accurate and complete understanding of the system;
- incorrect user expectations of the system;
- inappropriate, insufficient, or incorrect information being presented to the driver;
- a loss of concentration by the driver;
- the driver placing an overreliance on the system; and
- an incorrect assumption of how the user will interact with the design of the system.

Estimating the acceptability of the system's response to triggering conditions

To assess the system's response, the scenarios containing the triggering conditions must first be evaluated to establish whether the SOTIF is achievable in the first place.

- Known scenarios are covered by the verification activities to provide a final asessment of their acceptability.
- Assumptions that are considered during this evaluation can include the expected behaviors of the system, or the actions the user will presumably take.

The SOTIF is deemed achievable if the residual risk of the system causing a hazardous event is demonstrated as being lower than the acceptance criteria, and no known scenario is discovered that could lead to an unreasonable risk for the specified road users.

Managing the functional modifications that address SOTIF-related risks

Considerations for improving the SOTIF

The measures that address SOTIF-related risks are specified and then applied. As this work progresses, the information that is going into the specification and design must also be kept up to date. This work involves a number of considerations, including:

- **the specification and design** itself;
- **evaluating the risks** brought on by the hazardous behaviors;
- **identifying potential insufficiencies** in the specification, including performance insufficiencies and triggering conditions;
- **the verification and validation results** for both known scenarios and unknown hazardous scenarios; and
- **the arguments for releasing the SOTIF.**

Refining the system

The system is refined in an iterative process by considering the SOTIF measures. Then, the specification and design are updated with those measures, and then the risk of the intended functionality is evaluated using the updated specification and design.

This refined system is then evaluated in the V&V phase. If the residual risk from a known hazardous scenario is determined to be unacceptable, or an unknown and hazardous scenario is identified where the risk is unacceptable, or the overall residual risk is determined to be unacceptable, then the process is repeated to refine the system further. Both avoidance and mitigation measures can be used to reduce the SOTIF-related risks.

Even carefully designed and implemented SOTIF measures might not produce the expected outcomes, and they could result in unintended consequences. The monitoring and review activities prescribed in ISO 21448:2022, Clause 13, are an indispensable part of the process to help ensure that the measures in the SOTIF remain effective.

Modifying the system

For the sake of this discussion, *refining* the system is an exercise in continuous improvement, whereas measures for *modifying* the system are intended to maintain the intended functionality as much as possible. A common cause for modifying the system might be changes or improvements to the technology utilized by the system.

Modifications might include:

- increased performance and/or accuracy of the sensors;
- increased performance and/or accuracy of the actuators;
- increased performance and/or accuracy of the recognition and decision algorithms; and
- increasing the conspicuousness of the vehicle to improve the control performance of other traffic participants in response to hazardous behavior from the vehicle.

Functional restrictions

The purpose of restricting function is to maintain a partial functionality of a system by purposefully degrading the intended functionality. These measures might include:

- **Restricting intended functionality in specific use cases.** EXAMPLE: Having the vehicle go into a derate “limp home” mode when an obstructed catalytic converter is detected, allowing the vehicle to get the driver safely to their destination without allowing elevated exhaust temperatures that could further damage the catalytic converter or other elements in the exhaust system.
- **Removal of authority for the intended functionality in specific use cases.** EXAMPLE: Asking the driver to retake control of the vehicle when the sensors become blinded by a snowstorm.

Handing over authority from the vehicle to the driver

The purpose of handing over authority from a system to the driver, is to increase the controllability of the vehicle as the levels of driving automation become lower. This might be accomplished by modifying the human-machine interface with new instructions to the driver and additional information, modifying existing notifications, or by altering the fallback strategies.

In some instances, a handover might not be possible. If it is possible, the transition must be controllable and not generate additional risk. In other instances, it might also be advisable to incorporate some sort of timeout strategy that reduces the vehicle speed to zero in a safe manner if certain conditions are not met within a specified timeframe.

When designing solutions to these situations, it is advisable to refer to the practices for designing and evaluating ADAS.

Effective strategies for addressing reasonably foreseeable misuse

Customer education, in the form of information and training and their supporting materials such as training courses, marketing and sales presentations, and user manuals, are effective deterrents to reasonably foreseeable misuse by persons whose intent is to use the vehicle in a safe and responsible manner.

These materials should focus on the proper procedures for correct operation, emphasizing what should be done more so than what shouldn't be done. People tend to do what they read, and the processing of the "Should I or shouldn't I?" question can become mentally fatiguing if it has to be deciphered for every scenario.

Instead, most of the time emphasize what should be done. Reserve "don't do this" statements for sufficiently urgent caution and warning statements that are linked to the appropriate graphic iconography, at the decision juncture where there is the greatest risk if the driver makes the wrong choice. Typically, graphical icons can be deciphered much quicker than written words, if they are designed and presented well.

In an urgent situation where fractions of a second matter, you want the user to remember seeing examples of the right thing to do, rather than having to mentally sift through a catalog of wrong examples to try to flip their meaning on the fly to extrapolate the correct course of action.

The verification and validation strategy

The scope and purpose of verification and validation activity

The verification and validation strategy for the SOTIF provides an argument that the objectives have been achieved and demonstrates how the validation targets have been met. It must consider:

- the necessary evaluation of scenarios that are potentially hazardous;
- thorough and complete coverage of the relevant scenario space;
- necessary evidence in the form of analysis results, dedicated investigations, and test results; and
- the procedures neccesary to generate the evidence.

The rationale must be provided for the suitability of the selected V&V methods and the validation targets. This is foundation information that the remainder of the V&V process builds upon, and it is important to document it. While they may be obvious to the author, they might not be obvious to the other team members.

ISO 21448:2022, Clause 9.2, contains an extensive and detailed list of the information that should and can be considered.

In general, the V&V strategy focuses not only on the performance evaluation and risk identification within the ODD, but also on the boundaries that define the ODD, as well as other considerations outside the ODD. And, the V&V strategy must include verification that the system is not susceptible to outside interference or engageable from anywhere outside the ODD.

There is also a focus on the transitions from inside to outside the ODD accomplished by either escalation to the driver, or hand-off to the fallback system, whichever course is most appropriate for achieving the condition with minimal risk. This is what is referenced when arguing that sufficient coverage of the scenario space has been achieved.

Integration and testing

The verification and validation strategy covers the whole intended functionality of the vehicle. This Coverage including both the E/E elements and the elements of other technologies that are considered to be relevant to accomplishing the SOTIF. The verification and validation strategy also supports the monitoring of data from external sources that have been demonstrated to be relevant to the SOTIF in order to ensure that the system can't be engaged from anywhere outside the ODD. The validation targets provide evidence that the acceptance criteria are met.

A rationale for each defined effort must be provided, typically consisting of:

- the number or distribution of the scenarios;
- the number of experiments; or
- the duration of the simulation.

When determining the target values and the duration of the validation, if only a subset of scenarios proves to be relevant, then the exposure to that subset can be considered.

Evaluating the known and unknown scenarios

Evaluating known scenarios

The evaluation of known scenarios helps to determine whether the potentially hazardous scenarios are actually hazardous or not. Considerations include:

- The functionality of the system must behave as specified.
- Potentially hazardous behavior that is the result of specified behavior at the vehicle level, is evaluated to determine if it is appropriate and sufficient.
- The V&V strategy is evaluated to ensure that the known scenarios are adequately covered.
- The verification results must demonstrate that the validation targets have been met.

Issues can be assigned to different verification activities as appropriate. There are specific processes tailored to verifying the planning algorithm, the actuators, and integrated systems. There is also an evaluation process that addresses the residual risk due to known hazardous scenarios.

Evaluating unknown scenarios

The evaluation of unknown scenarios demonstrates that the residual risk from unknown hazardous scenarios fulfills the acceptance criteria with a sufficient degree of certainty. Considerations include:

- The specification and design;
- identifying the potential insufficiencies of the specification, and the performance insufficiencies and triggering conditions;
- the measures that address the SOTIF-related risks;
- the definition of the verification and validation strategy; and
- verification and validation results that demonstrate that the intended functionality behaves as expected in the known scenarios.

Unknown scenarios can be encountered in real-life situations. Methods to evaluate the residual risk arising from real-life situations that could trigger a hazardous behavior of the system when integrated in the vehicle, can be applied as illustrated in ISO 21448:2022, Table 11.

Testing in public areas

When conducting tests in public areas, it is possible that additional safety measures might become necessary to prevent or reduce the potential risk that might be posed to the public by the test vehicles, such as engagement of the emergency stop mechanism.

The selected methods are determined to be adequate for identifying potentially hazardous scenarios in Area 3, by using inputs that are representative for the use case as well as by focusing on rare or challenging environments in which to operate, as well as specific use cases, scenes or scenarios. A rationale must be provided that explains the adequacy of the methods that were selected.

Length and method of testing

Determining the appropriate length of the vehicle test must take into account the knowledge gleaned from previous vehicle programs, driver controllability, and the critical nauture of the routes that were selected.

When utilizing randomized input tests that include the injection of errors, the number of simulated scenarios can be chosen to correlate with a required test length and content that is representative of the geography found in the target market.

When considering the method of testing to be used such as test track scenarios, computer-based simulation, or open road driving, an appropriate allotment of kilometers to be driven or hours of operation is assessed and assigned to each method of testing. This allotment can and should be justified in the SOTIF documentation, so that others may understand how those numbers were determined.

Evaluating the achievement of the SOTIF

SOTIF is not created and published by individuals working in isolation. The work products that result from the SOTIF activities shall be reviewed for correctness, completeness, and consistency. The arguments for the achievement of the SOTIF shall be presented to the team and thoroughly evaluated. Then, a recommendation for approval or rejection of the release of the SOTIF shall be provided.

Methods and criteria for evaluating the SOTIF

To evaluate the SOTIF, important questions are considered.

- In all specified use cases, modifications to the design must reduce the risk sufficiently in accordance with the acceptance criteria. Have the hazards, potential functional insufficiencies, and triggering conditions been properly analyzed, and has any design modifications that are necessary to achieve the SOTIF been implemented and evaluated?

- Does the intended functionality accomplish a minimal condition of risk, when necessary?
- Does the intended functionality provide to the occupants or other road users a state that does not contain unreasonable risk, which also takes into consideration:
  - the specified interventions of the driver;
  - any reasonably foreseeable misuse;
  - warnings that have been specified for the vehicle occupants;
  - warnings that have been specified for the other road users;
  - the specified and purposeful degradation of the functionality; and
  - the DDT fallback needed to achieve the minimal condition of risk?
- Does the verification and validation strategy cover all the known hazardous scenarios?
- Does the verification and validation strategy provide an argument that the residual risk from unknown hazardous scenarios confidently meets the acceptance criteria?
  - Do the test results encompass the identified triggering conditions, making sure to cover all of the environmental conditions as well as both direct and indirect misuse?
  - Are sufficient activities included in the verification and validation strategy to limit the risk presented by both the known and unknown scenarios?
- Has a sufficient verification and validation process been completed and are the validation targets met, for the team to have a high degree of confidence that there is no unreasonable residual risk?
  - Has the intended functionality been exercised adequately enough to accurately and completely assess both the nominal behavior and the potentially hazardous behavior?
  - In the event of behavior that is determined to be hazardous, was evidence provided to argue the absence of unreasonable risk?

    - Did the testing provide sufficient coverage to support arguments for the completeness and robustness of the driving policy across all use cases and/or the ODD and the OEDR?

- Are the necessary means for realizing the operation phase activities (according to Clause 13) available?

Answering these questions might result in the need for additional steps. For example, if the activities in the operation phase that are described in Clause 13 have led to the creatioin or modification of any SOTIF measures, those measures are reviewed in Clause 12. And, the examination of the results from the SOTIF activities can also be jointly considered with the ISO 26262-2 functional safety assessment.

Recommendation for the release of the SOTIF

Based on evidence of the methodology from 12.3, a recommendation of “acceptance”, “conditional acceptance” or “rejection” for release can be determined. In case of “conditional acceptance”, the conditions are documented, and their fulfilment is verified before final release.

The evaluation of the achievement of the SOTIF is then considered to be documented.

Operation phase activities

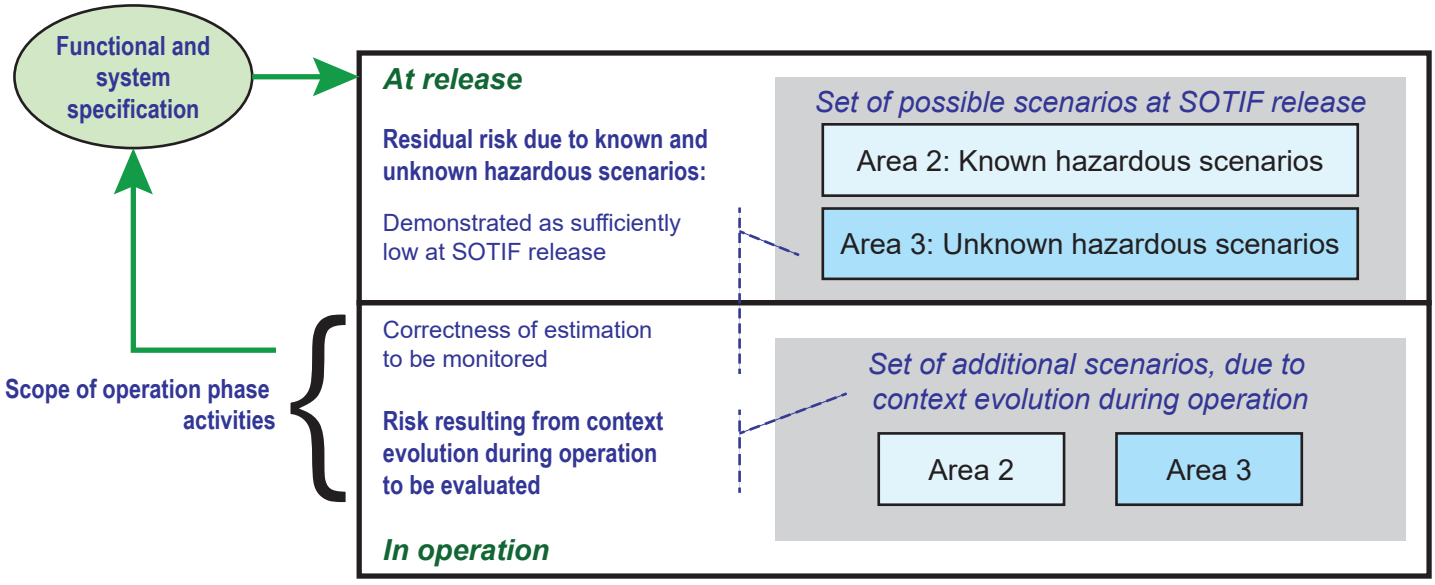
The purpose of this clause is to achieve these objectives:

- before the release of the SOTIF, define a process for monitoring field activities to ensure achieving the SOTIF during the operation phase; and
- perform the field monitoring process in a manner that maintains the achievement of the SOTIF throughout the duration of the operation phase.

The SOTIF activities described in Clauses 5 through 12 strive to reduce the risk to an acceptable level at the time that the SOTIF is released. However, there are conditions under which that risk evaluation might be reconsidered. For instance:

- if the operation of the functionality in the field uncovers a previously unidentified hazard;
- if the operation of the functionality in the field uncovers a previously unidentified insufficiency and/or triggering condition; and
- if assumed parameters such as environment conditions or traffic regulation change, compared to how they were defined during the development of the functionality.

Derived from ISO 21448:2022, Clause 13.2, Figure 16 — Scope of operation phase activity[1]



Unsolved challenges

The challenge of managing updates

In adopting these responsibilities, vehicle systems that are released out on the road must be updated over time. Every partner in the value chain will have to be able to validate that the system they have just updated is now safer than it was before the update, regardless of when it was first built. And they will have to do that every time an update is issued, resulting in an almost constant revalidation loop. This will require significant involvement from every vendor and manufacturer that impacts the vehicle, especially the software and semiconductor suppliers.

What the future may allow is over-the-air updates of software. But software updates will need to undergo the same Verification and Validation (V&V) rigor as the original release (also known as “regression testing”). This is where an automated scenario execution will pay dividends. Also, AI might be employed in the training realm, allowing values to be locked in for production.

Retroactive compatibility

Will the requirement for retroactive compatibility make sign-off harder or easier as you progress through the design cycle?

That is a tough question to answer because it is a system-level problem. Companies are going to have to perform validation in a way that exercises everything in that system properly, including the semiconductors. The industry has not figured out how to do that within those complex systems yet, although remote automated around-the-clock simulation is likely to play a very important role, with the most prominent reason being the sheer bandwidth required to test all the scenarios in a timely and cost-effective manner.

Updating from lessons learned

Most people are not used to thinking of their cars as evolving systems. Instead, they are used to thinking of them as systems that degrade over time. Traditional vehicles are bought, they are used, they break down and are repaired. They rust, and their performance steadily degrades over time. Eventually, they wear out and are scrapped.

In comparison, a modern autonomous vehicle must be kept at peak operating performance for its entire lifetime. So, as issues arise, are they addressed inside the chips, circuits, and mechanicals, or externally? The answer is: all the above. They are addressed as a whole system, at any of those subsystem blocks, or within the components that comprise those subsystems.

The variety in models and available features adds considerable complexity and difficulty. Original Equipment Manufacturers (OEMs) should be working constantly to find all the unknown/unsafe scenarios in order to move them to a known/unsafe state. This work is performed during the research and development phase, or the production development phase, regardless of whether AI is used in the development process or not. Then, the learned values are locked in for production.

As time passes, the algorithms become “smarter” as they learn new scenarios and better ways to respond, but this improvement work is still done in a development environment. When an iterations is successfully completed, then the fielded systems can be updated, which in turn feeds into the need for further refinements and updates.

And after every update, there must be complete and thorough validation for safety.

The importance of accurate and complete data

Frankly, we need much better diagnostics on these vehicles that we have today. And, when there is a problem, we are going to need high-quality forensics. Accurate and complete data is paramount and makes everything else possible.

Diagnostics is going to be an area of significant change. The diagnostics must provide a better and more comprehensive understanding of how the system is reacting to the world around the vehicle than the industry is capturing right now. We must develop the ability to take that information, send it back to a data center, process it thoroughly, and perhaps identify new learning to help develop future software updates.

In instances where the car crashes, forensics will be required to understand what the system was thinking before the crash happened, and what its view of reality was before it did whatever it did. That context is critical to achieving an accurate and complete understanding of what actually happened, and it is going to require work in both the software and hardware realms to get access to all that data. And it is going to take time and the development of appropriate validation processes to ensure that the forensic data is being interpreted and applied accurately.

We have our work cut out for us.

Summary

Each SOTIF scenario can be sorted into one of four classifications: *known hazardous, known not hazardous, unknown hazardous, and unknown not hazardous.* These classifications indicate the nature and risk within their scenarios and provide an approachable starting point.

The first priority is to cover the known scenarios, the things that we already know, namely, the known not hazardous and the known hazardous. Those scenarios are pretty straightforward. In comparison, testing the unknown becomes a more abstract exercise.

The unknown hazardous is the nightmare scenario, where you don't know what you don't know, and you have no visibility of exactly how much risk there is. It is addressed in part using testing simulations such as hardware-in-the-loop (HIL) systems that allow for testing a significant variety and volume of scenarios in rapid order, in a safe and cost-effective manner.

The unknown not hazardous scenarios can offer surprises, but a rugged system design can absorb most of the impacts resulting from an unknown yet not hazardous scenario.

Driving conditions vary depending on road conditions and the environment. Statistical distributions are employed to reflect how we expect the vehicle to be operated under these different types of conditions. In these, we err towards testing more on the side of the less safe. This helps to compensate for the reality that the industry does not yet have a concrete plan for systematically injecting randomness into this work. We as an industry have made significant progress in systematically managing these challenges, with SOTIF playing a pivotal role. But there is still much to be done to create testing that reflects the randomness of real life.

Software is the lifeblood of safety systems. It is instruction and learning, the source and the recipient of data. And frankly, it is a world where best practices are known but not always applied. Software has played a pivotal role in wonderful solutions and terrible missteps. And as software inherits an increasingly important role, the demands of SOTIF are challenging software developers to raise the bar on the quality of their code. Software, more than any other element, has the potential to raise the quality of the entire system.

The automotive industry in general is still trying to wrap its arms around the challenge of finding the optimal way of integrating SOTIF organically into the design process from the start. The standards certainly help, but the task of integrating SOTIF can be eased somewhat by reinforcing the importance of quality data. The automotive realm has seen the consequences of allowing quality to slip. **Quality data acted upon with discipline is the foundation upon which SOTIF success is built.** Our functionally safe future depends on it.

Bibliography

[1] ISO 21448:2022, *Road vehicles — Safety of the intended functionality.*

[2] *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles,* SAE Recommended Practice J3016_201806, https://www.sae.org/standards/content/j3016_201806

[3] *COMMISSION RECOMMENDATION of 22 December 2006 on safe and efficient in-vehicle information and communication systems: update of the European Statement of Principles on human machine interface (2007/78/EC)*: https://data.europa.eu/eli/reco/2007/78/oj