IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2025.0000000

Tailoring STPA for SOTIF: Terminology Mapping and Methodological Extension

ALEXANDER AHLBRECHT[1]†, NAYEL FABIAN SALEM[2]†, LINA PUTZE[3]†, INGO STIERAND[3]†, UMUT DURAK[1], MARCUS NOLTE[4], ECKARD BÖDE[3]

[1]German Aerospace Center, Institute of Flight Systems, 38108 Braunschweig, Germany
[2]Technical University of Braunschweig, Institute of Control Engineering, 38106 Braunschweig, Germany
[3]German Aerospace Center, Institute of Systems Engineering for Future Mobility, 26121 Oldenburg, Germany
[4]KTH Royal Institute of Technology, Dept. of Engineering Design, Unit Mechatronics, 10044 Stockholm, Sweden

Corresponding author: Alexander Ahlbrecht (e-mail: alexander.ahlbrecht@dlr.de).

†These authors contributed equally to this work, ''This work was supported by the V&V4NGC project of the German Aerospace Center''

ABSTRACT According to ISO 21448, it is essential to consider the Safety of the Intended Functionality (SOTIF) to ensure the safety of automated vehicles. A key objective for SOTIF is the identification and analysis of triggering conditions and functional insufficiencies. To support this objective, ISO 21448 suggests the System Theoretic Process Analysis (STPA) as a suitable analysis technique. Although STPA is a promising hazard analysis method, it was not specifically developed for SOTIF. Consequently, it is necessary to create a terminology mapping and methodological extension in order to adapt STPA for SOTIF. For example, STPA terms such as ''loss'' have more specific meanings than their ISO 21448 counterparts. At the same time, SOTIF requires a systematic analysis of scenarios to identify triggering conditions and functional insufficiencies. Although STPA is suitable for scenario-based analyses, it does not guide the scenario specification. To address the identified gaps, this article proposes the use of SOTIF-specific terminology mapping and an extension to the STPA method. These extensions include a behavior specification and hazard identification approach, building the foundation for a STPA tailored for SOTIF. With these changes, it becomes possible to trace the STPA artifacts to ISO 21448 objectives.

INDEX TERMS ISO 21448, SOTIF, STPA, behavior specification, hazard identification

I. INTRODUCTION

ASSURING safety for automated vehicles comes with a number of challenges [1] that must be addressed systematically. New methods and standards must be developed, as discussed by [2] and [3]. In particular, automated systems require not only the assurance of functional safety. It also needs to be ensured that the specified functionality of the system is inherently safe. Accordingly, the ISO 21448 standard gives development guidance with the focus on vehicles' *Safety Of The Intended Functionality* (SOTIF) [4]. In its appendix, ISO 21448 mentions *System Theoretic Process Analysis* (STPA) [5] as one of the methods to conduct SOTIF analyses. Although the application of STPA has already shown promising results in supporting SOTIF activities [4], [6], [7], it was not specifically designed for ISO 21448 compliance.[1] Resulting gaps are highlighted by [11], [12].

For instance, ISO 21448 requires the identification of *functional insufficiencies* and *triggering conditions* that can cause *hazardous behavior* of the vehicle. To identify functional insufficiencies and triggering conditions, detailed knowledge of the operational context of the systems is required. STPA can incorporate information on operating conditions, as shown in [7], but does not provide a systematic approach to their specification [11]. In addition to methodological gaps, a major obstacle for the integration of STPA into SOTIF is the difference in terminology. The problems involved include mismatches in definitions of similar terms, as well as the fact that the concepts used in ISO 21448 and STPA do not necessarily cover the same aspects.

This article contributes to the above challenges in the following ways. First, we provide a consistent alignment of the STPA and SOTIF terms. This sets the foundation for the second contribution, the adaptation of STPA for the identification of triggering conditions and functional insufficiencies within the scope of ISO 21448. To that end, we propose an integrated approach that combines STPA with a method

[1]As STPA is designed to support system safety [5], it can also be applied to security aspects [8]. Related tasks are, for example, specified in ISO/SAE 21434 [9]. However, this article focuses on the scope of SOTIF as specified in ISO 21448 [4]. The integration of safety and security aspects is, for example, specified in ISO/TS 5083 [10].

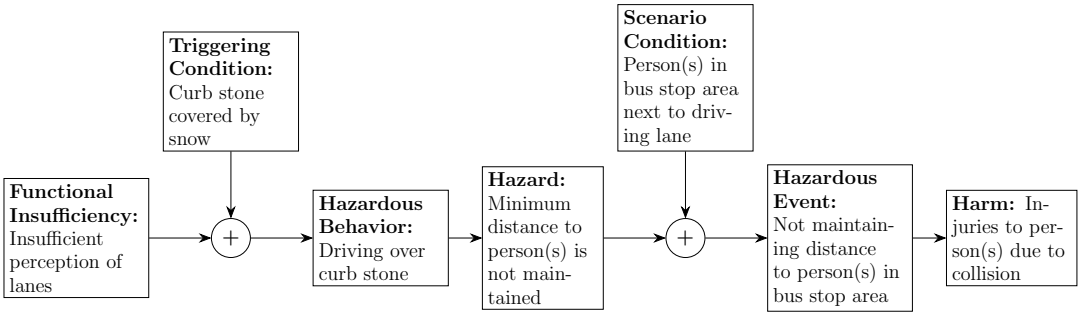


FIGURE 1. **Hazardous event model of ISO21448 as discussed in [13]**

TABLE 1. **SOTIF-related terminology based on standards and related publications**

Term	Definition
Harm	physical injury or damage to the health of persons. ([14], 3.74)
Hazard	potential source of harm. ([15], 3.2)
Hazardous Event	event that is a combination of a hazard and a scenario containing conditions in which the hazard can lead to harm. [13]
Hazardous Behavior	behavior of the vehicle leading to one or more vehicle level hazards. [11], [13]
Scenario	description of the temporal relationship between several scenes in a sequence of scenes, with goals and values within a specified situation, influenced by actions and events. ([4], 3.26 based on [16])
Use Case	description of a suite of related scenarios. ([4], 3.32)
Scenario Condition	operating condition in a scenario in which the hazard can lead to harm.
Triggering Condition	specific condition of a scenario that serves as an initiator for a subsequent system reaction contributing to either a hazardous behavior or an inability to prevent or detect and mitigate a reasonably foreseeable indirect misuse. ([4], 3.30)
Functional Insufficiency	performance insufficiency or insufficiency of specification. ([4], 3.8)
Performance Insufficiency	limitation of the technical capability contributing to a hazardous behavior or inability to prevent or detect and mitigate foreseeable or indirect misuse when activated by one or more triggering conditions. ([4], 3.22)
Insufficiency of Specification	specification, possibly incomplete, contributing to either a hazardous behavior or an inability to prevent or detect and mitigate a reasonably foreseeable indirect misuse when activated by one or more triggering conditions. ([4], 3.12)

for the specification of intended behavior and a method for identification of hazardous behavior. To facilitate compliance with ISO 21448, we provide a mapping that links the results of the integrated approach to ISO 21448 objectives. In summary, this article addresses terminological and methodological gaps when applying STPA for SOTIF.

In order to showcase these contributions, this article is structured as follows. Section II introduces the main concepts of SOTIF and STPA. Then, we summarize the related work for the integration of SOTIF and STPA, highlight the gaps between STPA and SOTIF, and obtain requirements based on this analysis in Section III. To address the identified gaps, we map the terminology between ISO 21448 and STPA in Section IV and elaborate in Section V on a methodological extension of STPA. In Section VI we demonstrate the application of the integrated approach by analyzing a people mover concept. Finally, we discuss results and limitations in Section VII and provide a conclusion in Section VIII.

II. BACKGROUND

A. SAFETY OF THE INTENDED FUNCTIONALITY

In contrast to functional safety, which investigates system failures and their consequences, SOTIF is concerned with the safety of the functionality itself. SOTIF covers aspects such as gaps in the specification, limited system performance, and the inability to detect or prevent reasonably foreseeable misuse. ISO 21448 [4] gives guidance how to address SOTIF for road vehicles with driving automation systems, that is, SAE level one to five [18]. The standard provides a broad framework that structures essential SOTIF activities and defines overarching objectives. For certain activities, it also suggests specific methods or criteria that may be appropriate to support SOTIF.

Fig. 1 illustrates the *hazardous event model*, which constitutes the foundation for the activities specified by the standard to ensure SOTIF. The general idea is that the system specification and design may contain so-called *functional insufficiencies* – limitations of either the specification or the system performance. Functional insufficiencies can be activated by one or more *triggering conditions* during runtime, resulting in a certain *(hazardous) behavior* of the vehicle, which – given some additional *scenario* conditions – can lead to *harm*. It should be noted that the harm covered by the standard is limited to physical injuries. Detailed definitions of important ISO 21448 terms are listed in Table 1 and are supplemented by related work where needed.

The main goal of ISO 21448 is to achieve SOTIF by a systematic analysis and treatment of the individual artifacts in the hazardous event model that contribute to the probability of occurrence of harm and the severity of that harm (that is, risk). To achieve this goal, ISO 21448 formulates several objectives for the system development process (specification, design,

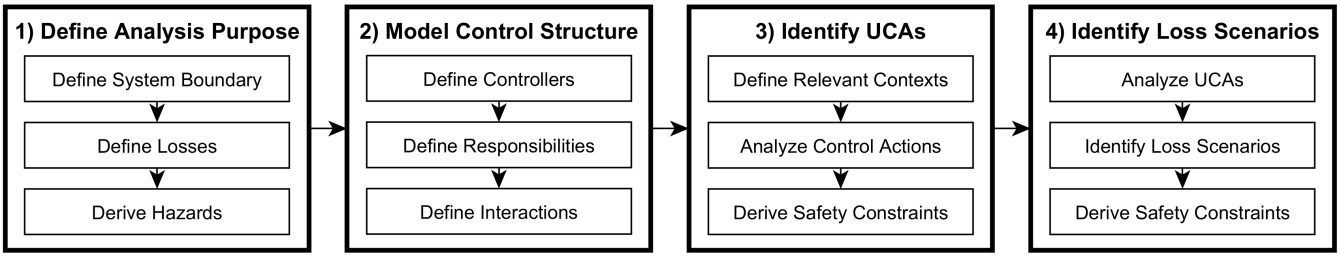


FIGURE 2. **Analysis workflow of STPA**

TABLE 2. **STPA terminology [5], [17]**

Term	Definition
Loss	The act or fact of being unable to keep or maintain something valued by a stakeholder.
Hazard	System state or set of conditions that, together with a particular set of worst-case environmental conditions, will lead to a loss.
Control Action	A command, instruction, directive, or other information provided by a controller to control a process and achieve goals.
Feedback	A value, measurement or other information provided to a controller to achieve its goals.
UCA	A control action that, in a particular context and worst-case environment, will lead to a hazard.
UCA Context	A state or condition that would make the control action unsafe
Loss Scenario	The causal scenarios or factors that can lead to unsafe control actions and to hazards.
Causal Scenario	The reasons, situations, or incidents that allow a control action to occur.
Causal Factor	A circumstance, fact, or influence contributing to a causal scenario.
Process Model	A representation of a controller's internal beliefs used to make decisions. Process models may include beliefs about the controlled process or other relevant aspects of the system or the environment.
Mental Model	An explanation of someone's thought process about how something works in the real world. It represents the surrounding world, the relationships between its various parts, and a person's intuitive perception about their own acts and consequences.

verification, and validation) and the operation phase, which are centered around the hazardous event model and require a rigorous description of the operational context [19]. The objectives relevant to the present work can be summarized as follows:

Obj. 1. The specification and design of the system and its operational context shall be sufficient to carry out SOTIF-related activities and shall be updated iteratively in response to these activities ([4], Clause 5.1).

Obj. 2. Hazards at vehicle level that arise due to insufficiencies of intended functionality shall be identified ([4], Clause 6.1).

Obj. 3. Risks caused by hazardous behavior and the corresponding scenarios in which hazardous behavior can lead to harm, including concrete parameters, must be systematically identified and evaluated ([4], Clause 6.1).

Obj. 4. Potential functional insufficiencies and triggering conditions contributing to hazardous behavior shall be identified and their impact on SOTIF shall be evaluated ([4], Clause 7.1).

B. SYSTEM THEORETIC PROCESS ANALYSIS

The SOTIF related objectives need to be addressed with a systematic hazard analysis method. ISO 21448 mentions STPA as a suitable method to identify functional insufficiencies and triggering conditions [4]. STPA is a top-down general purpose hazard analysis method [20]. It consists of four main steps that are detailed in the STPA handbook [5] and shown in Fig. 2.

First, the purpose of the analysis is defined by specifying the system boundary as well losses that shall be prevented. In addition, hazards are derived that can cause the losses. Second, the system under analysis is represented in a control oriented perspective. Therefore, the system is modeled as a hierarchic control structure connecting controllers, actuators, sensors, and the controlled process. Between control structure elements, interactions are visualized as control actions (commands) and feedback interactions (information exchange). To identify potential sources of hazards in the control structure, control actions are analyzed. If there is a context where a control action can be dangerous, it is documented as *Unsafe Control Action* (UCA). Finally, in the fourth step, the causal reasons for UCAs are explored and documented in the form of loss scenarios. To identify loss scenarios, causal factors such as system interactions, control algorithms, and related assumptions (e.g., process or mental models) are investigated. To ease the identification of issues related to process/mental models during the analysis, corresponding assumptions should be documented beforehand. Throughout the process, mitigating safety requirements are derived and documented to address potentially unsafe behavior and its causal factors. In addition to the process description, STPA's terminology is shown in Table 2.

Considering the focus of the method, it becomes clear why STPA is a promising candidate to address the SOTIF objectives. STPA not only focuses on component failures, but tries to identify inadequate system interactions and specification

issues including socio-technical and organizational aspects. However, since it is not specifically designed for SOTIF, some terminological and methodological gaps exist and are highlighted in the related work.

III. RELATED WORK AND RESULTING REQUIREMENTS

There is significant interest in applying STPA to support SOTIF. This is reflected not only by the fact that STPA is explicitly mentioned in ISO 21448 [4] but also by the amount of existing related work. The following discussion focuses on identifying strengths and weaknesses of the application of STPA for SOTIF. Based on this analysis, we elicit requirements for a clear connection of STPA and SOTIF concepts.

A. RELATED WORK

Multiple publications apply STPA for SOTIF relevant systems and subsystems. In [21], STPA is utilized to analyze an autonomous perception system. Similarly, [22] analyzes an intelligent railway driving assistance system and [6] investigates a machine learning-based perception system with STPA. Other work focuses on the utilization of STPA for specific SOTIF activities. An example is [7], which describes how to utilize STPA to support the creation of test scenarios for SOTIF. This relation between STPA and SOTIF related testing is also highlighted in [23].

However, some articles also raise questions about the applicability of STPA to SOTIF. The authors of [12] argue that STPA needs more guidance and detail for the control structure to cover perception-related risk causes that are within the scope of ISO 21448. [11] analyzes the potential support of STPA in the identification of SOTIF triggering conditions. The authors agree that the general concepts in STPA and SOTIF are compatible (e.g., hazards, harm). However, despite being a fundamental concept of the SOTIF analysis according to ISO 21448, the concept of hazardous behavior is neither defined in ISO 21448 nor in the context of STPA. In contrast to the application example within the appendix of ISO 21448, the authors of [11] argue that the concept of UCAs is not always synonymous with the term hazardous behavior. Their rationale is that hazardous behavior often refers to vehicle behavior instead of system interactions. Due to the importance of hazardous behavior for SOTIF, [11] highlights that more work is needed to provide guidelines for a more rigorous identification of hazards, hazardous behavior, and the context in which it is exhibited. Moreover, to apply STPA within the SOTIF context, the authors propose the addition of a dedicated step to map STPA results to SOTIF triggering conditions. In summary, a terminology mapping is needed (Req. 1) and a more systematic approach to identify hazardous behavior (Req. 2) would be required to tailor STPA to SOTIF.

Since the consideration of scenarios and the related system context is important not only for SOTIF but also to perform STPA, some literature discusses how to specify the context within STPA. For example, [24] shows how the context influences the analysis results and discusses how a variety of context variables (operational and system states) can be integrated into an STPA analysis. However, how context variables can be systematically derived is not covered in detail. In [25], an analysis of the *Operational Design Domain* (ODD) is performed to derive critical scenarios that are then analyzed using STPA. The integration of ODD and STPA is further refined and applied to autonomous systems in the maritime domain in [26]. These publications show that STPA benefits from a systematic definition of the operational context to identify SOTIF related issues (Req. 3). Although not in the scope of this article, structured ODD models can be used as a basis for the specification of operational contexts. One example is the 6-layer model discussed in [27], [28].

B. REQUIREMENTS TO TAILOR STPA FOR SOTIF

We summarize the findings of related work as follows. The interest in applying STPA for SOTIF is high and is even suggested in the ISO 21448 [4]. Although some related work shows the application of STPA for SOTIF systems, compliance with concrete ISO 21448 objectives is currently not emphasized. Other related work applies STPA for specific SOTIF activities such as testing. At the same time, we identified gaps in the literature that could be addressed to improve the applicability of STPA for SOTIF. More specifically, the goal of this article is to tailor STPA to the application of the SOTIF standard ISO 21448. As indicated above, a particularly important element in this regard is to enable STPA to identify triggering conditions and functional insufficiencies. Based on this goal and the lessons learned from the related work we formulate requirements for this article that we discuss in the following.

Req. 1. STPA terminology shall be mapped to the SOTIF hazardous event model. To apply STPA to identify functional insufficiencies and triggering conditions, the alignment of the concepts and terminology related to the hazardous event model in Fig. 1 is important. Without a systematic alignment of the concepts and terminology between STPA and SOTIF, the analysis might be guided in a different direction than what is required to address SOTIF related issues. At the same time, an alignment is needed to identify which SOTIF artifacts are covered by STPA and which have to be provided by complementary analyses. Hence, a dedicated mapping between STPA terminology and SOTIF terminology is needed to allow a shared foundation for the analysis. This requirement is addressed in Section IV.

Req. 2. STPA shall be complemented with a systematic process to identify hazardous behavior. In addition to the terminology gap, the identified methodological gaps from the related work must be covered. An important aspect concerns the identification of hazards in the STPA process. The application of STPA to SOTIF would benefit from more detailed guidance on how to specifically identify hazardous behavior and respective hazards. Since hazard identification can benefit from a systematic identification of hazardous behavior, a complementary process is required to identify SOTIF-related

TABLE 3. Mapping STPA and SOTIF terminology with an example

SOTIF Terminology	Example	STPA Terminology
Harm	A collision with physical injuries is caused by	Loss
Hazard	a violation of the minimum distance between a vehicle and a pedestrian	Hazard
Scenario Condition	during a stop at a bus stop where a person is waiting.	UCA Context
Hazardous Behavior	The collision occurs because the vehicle does not keep its lane and drives over a curb stone.	Unsafe Control Action
Triggering Condition	The reason is that the curb stone is covered by snow	Causal Factors
Functional Insufficiency	while the vehicle was not designed to handle covered curb stones.	

hazardous behavior. This requirement is addressed in Section V-B.

Req. 3. STPA shall be complemented with a systematic process to define operating conditions. Another topic identified by related work is the fact that the detail of the analyzed operating conditions heavily influences STPA's results. Although STPA can include operating conditions, it does not provide a systematic approach to define them. Therefore, a complementary process for defining and integrating operating conditions shall be provided. This requirement is addressed in Section V-A.

IV. TERMINOLOGY MAPPING FROM STPA TO SOTIF

Starting with Req. 1, STPA's terminology must be aligned with SOTIF's hazardous event model to create a shared foundation for a SOTIF-tailored STPA application. Table 1 and Table 2 highlight important terms of ISO 21448 and STPA, respectively. Although there are some similarities in the terminologies, ambiguity remains. This can complicate the mapping of STPA artifacts and SOTIF work products. For instance, the term hazard is defined within STPA's terminology as a ''system state that [...] could lead to a loss'', while the SOTIF hazard refers to the broader definition of ''a potential source of harm''. Another example that was criticized in related work is the postulated equivalence of STPA's UCA with hazardous behavior [11]. However, this equivalence is only valid when the UCAs are defined at the vehicle level. The differences between the terms and concepts must be considered when STPA is applied in the SOTIF context.

To provide a consistent terminology baseline, our proposed mapping is shown with an example from a current research project that will be introduced more thoroughly in Section VI. In a nutshell, for the research vehicle *Ushift* we consider a people mover use case, where persons are picked up and dropped off by an automated vehicle at regular bus stations. For the terminological mapping, we divide a potential accident description into different parts and map STPA and SOTIF terminology in Table 3.

Term 1. Harm: For SOTIF ISO 21448 defines harm as a physical injury or damage to the health of people. In contrast, STPA uses the broader concept of losses to define the purpose of the analysis. Losses can cover everything that is of value to stakeholders and therefore present a superset of harm. To execute STPA in the scope of ISO 21448, the set of STPA's losses can be restricted to the subset of physical injuries. While this is not explicitly included in ISO 21448, [29] and [30] argue that a broader interpretation of harm in the SOTIF context could lead to a more balanced safety assessment (especially considering ethical values).

Term 2. Hazard: Is used differently by ISO 21448 and STPA. In ISO 21448, a hazard is simply stated as a potential source of harm. In STPA, the term hazard is more specific and is defined as a system state that could lead to a loss. Since STPA's more precise definition is in line with ISO 21448's definition, the suggestion is to use STPA's definition for easier integration.

Term 3. Scenario Condition: ISO 21448 differentiates between scenario conditions that cause hazardous behavior at vehicle level – the triggering conditions – and scenario conditions in which the hazard can lead to harm. In STPA, the latter are captured in the form of the UCA context. The UCA context is a state or condition that would make the control action unsafe.

Term 4. Hazardous Behavior: While hazardous behavior is not explicitly defined within ISO 21448, assumptions about the term can be derived. For example, [11] defined the term as an action or movement at the vehicle level that causes one or more hazards. In the appendix example in ISO 21448 [4], the STPA term UCA is used as an equivalent to hazardous behavior. A note explains that a UCA always leads to one or more vehicle-level hazards. At the same time, it is important to mention that UCAs may cover any abstraction level from vehicle down to low-level interactions. As hazardous behavior is defined on the vehicle level, UCAs can be considered synonymous when they are also defined on the vehicle level.

Term 5. Triggering Condition: Is defined by the ISO 21448 as an initiator of hazardous behavior or inability to prevent, detect, mitigate reasonably foreseeable misuse. Triggering conditions focus on the operational environment conditions that need to be present in order to trigger an insufficiency. A related term in STPA is the causal factor which describes the reasons for a UCA. Causal factors focus on the causal impact on, for example, a control action or controller. In STPA, causal factors not only cover triggering conditions, but can

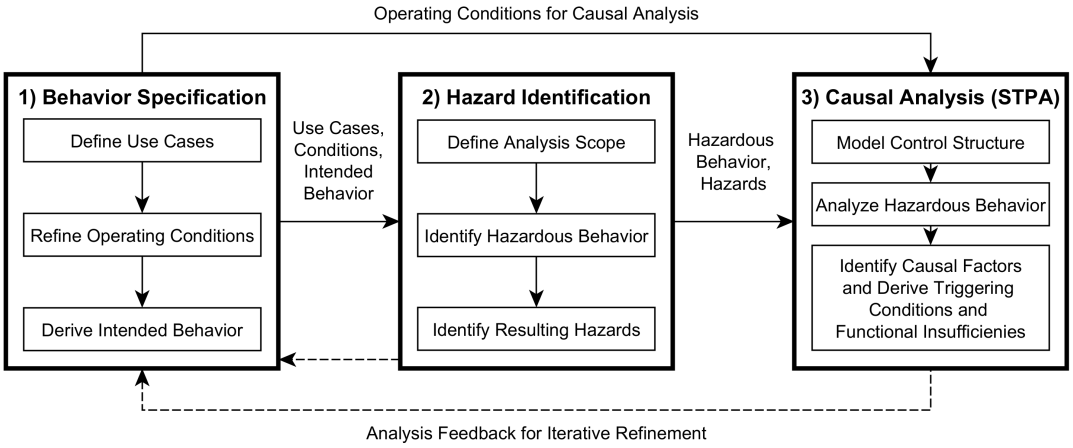


FIGURE 3. **An integrated approach for a SOTIF-tailored STPA in the scope of ISO 21448**

contain multiple types of causes. Thus, triggering conditions represent a specific subset of STPA's causal factors (e.g., factors related to interactions with the operational environment). As a result, a categorization of causal factors is needed which checks if the causal factor is a triggering condition.

Term 6. Functional Insufficiency: Is defined by the ISO 21448 as performance insufficiency or specification insufficiency. Functional insufficiencies focus on system properties related to the specification and technical capabilities that can lead to hazardous behavior if activated by one or more triggering conditions. A related term in STPA again is the causal factor that specifies the reasons for a UCA. Causal factors focus on the causal impact on, for example, a control action or controller. In STPA, causal factors not only cover functional insufficiencies, but can contain multiple types of causes. Thus, functional insufficiencies represent a specific subset of STPA's causal factors (e.g., process model related factors). Consequently, a categorization of causal factors is needed that checks if the causal factor represents a functional insufficiency.

V. METHODOLOGICAL EXTENSION TO STPA FOR SOTIF

While the terminology mapping suggested above allows us to interpret STPA in the SOTIF context, this section focuses on methodological extensions. To achieve a systematic identification of triggering conditions and functional insufficiencies and address the deficiencies identified in the related work, we integrate STPA with (1) a method specifying the vehicle's operational context (Req. 3), and (2) a method identifying hazards and hazardous behavior (Req. 2).

Fig. 3 shows the integration of the three methods. Initially, a *behavior specification* includes the definition of use cases, the refinement of operating conditions, and the derivation of the intended behavior of the automated vehicle. The results serve as input for a *hazard identification* procedure, which analyzes the deviations and effects of the intended behavior of the vehicle and the operating conditions to identify hazardous behavior and hazards. Hazardous behavior, the resulting hazards, and the operating conditions serve as input to STPA-based *causal analysis* to identify functional insufficiencies and triggering conditions. The results from the *hazard identification* and *causal analysis* can serve as feedback for an iterative refinement of the *behavior specification*. But, the iterative refinement will not be the focus of this article.

A. BEHAVIOR SPECIFICATION

As the first step of the integrated approach, we complement STPA with a behavior specification (see Fig 3). The goal of this integration is to systematically specify the operating conditions considered and the intended behavior of the automated vehicle under these conditions as specified in Req. 3.

This step employs the method proposed by [31]. It uses knowledge from sources such as the traffic code to specify the intended behavior under certain operating conditions. The intended behavior is specified in terms of driving maneuvers (see Table 4). To establish causal relations between operating conditions and driving maneuvers, the authors use directed acyclic graphs (see Fig. 4). The nodes in the graph are either atomic operating conditions, facts, or maneuvers. Operating conditions describe scene elements such as road infrastructure, traffic signs, or weather conditions. Facts capture inferred operating conditions such as a pedestrian's crossing intention or a valid signaling of a bus stop. These facts can be derived, for example, from the traffic code [31]. Maneuvers (according to [32]) can be distinguished as lateral maneuvers (e.g., keep lane, change lane) and longitudinal maneuvers (e.g., start, stop, keep target speed), where the combinations of both are captured in the behavior specification. The edges connecting the nodes capture the causality for the specified behavior as assumed during specification and design. The behavior specification does not attempt to model a decision-making logic for runtime algorithms but captures assumptions regarding the relevant operating conditions in which the specified maneuvers are considered as intended.

For the application of [31] in our integrated approach, we

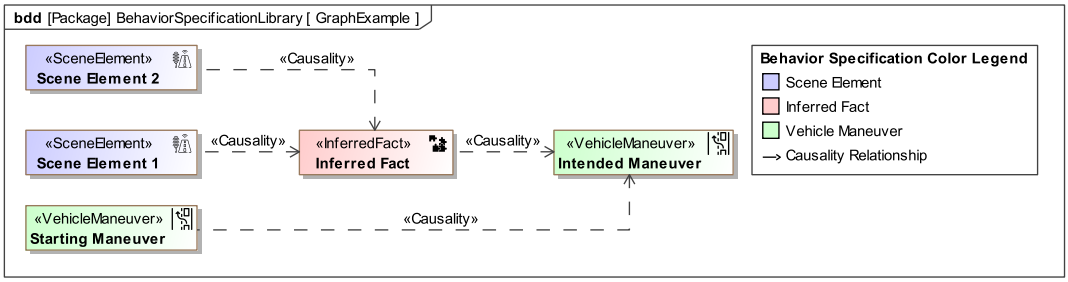


FIGURE 4. **Graph-based behavior specification**

TABLE 4. **Vehicle maneuvers based on [32]**

Longitudinal Maneuver	Lateral Maneuver
Start	Keep lane
Stop	Change lane
Follow Desired Speed	Pass
Follow Traffic Participant	

first set the scope of the behavior specification by defining a preliminary system boundary and relevant use cases. Subsequently, the operating conditions for the use cases are refined to provide sufficient detail for the hazard identification (as proposed in [33]) and causal analysis using STPA. The intended behavior is specified on the basis of these operating conditions. Specified maneuvers and operating conditions are used to identify hazardous behavior during hazard identification (Section V-B). Furthermore, the integration of STPA with a specification of the operational environment influences the design of the STPA control structure, the consistent specification of loss scenarios, and the identification of triggering conditions and functional insufficiencies during STPA (Section V-C).

Generally, other approaches to behavior specification can be used for this step as well. For example, meta-models for a systematic description of the operational environment are proposed in [34], [35], [36]. Furthermore, there are alternative formalisms to specify intended behavior [37]. For the integration into our approach, a necessary condition is that the behavior specification provides traceability between the intended behavior at the vehicle level and the operating conditions under which this behavior should be exhibited. In this regard, the description of external behavior using maneuvers as proposed by [32] can also be performed using the taxonomy provided by [38].

B. HAZARD IDENTIFICATION

The second step of the proposed integrated approach is the hazard identification. The objective of this step is to determine whether any vehicle behavior may lead to harm, addressing Req. 2. The identification of hazards provides the basis for a comprehensive safety argumentation. Consequently, a systematic approach must be employed that enables the argument that all foreseeable hazards are identified. Although the definition of hazards is part of STPA, it does not provide systematic measures to identify hazards as described by [11]. At the same time, a precise link of STPA results to vehicle-level hazardous behavior is needed to achieve ISO 21448 objectives. To address these shortcomings, we employ the keyword-based brainstorming approach proposed by [39] while integrating and facilitating the results of the behavior specification. The applied method is conceptually based on the HAZOP (Hazard and Operability Study) approach, which systematically identifies potential deviations through structured variation of conditions. Nevertheless, alternative systematic top-down techniques that can be applied at the system level may likewise be suitable to identify hazards and hazardous behaviors.

The general idea of the approach developed by [39] is to define abstract use cases that cover the target operational domain and to derive hazardous behaviors and hazards through a gradual refinement process. In the first step of the approach, each intended use case is combined with all the different maneuvers that the vehicle under consideration can perform (see Table 4). Next, for each maneuver, a context in the form of refined operating conditions within the considered use case is annotated in which the selection of the maneuver would be correct. In order to identify hazardous behaviors, each combination of use case, maneuver, and context is then combined with keywords such as "provided", "not provided", "less", "more" or "too early". These keywords can be applied to either the maneuver or the context to explore variations that may be hazardous. If this brainstorming process reveals any behavior of the system that could cause harm within the considered use case, this hazardous behavior and all subsequent effects are documented. These effects include observable effects in the scenario, hazardous events, hazards, and additional scenario conditions.

In the context of the integrated approach, the behavior specification of Section V-A already provides a systematic way to derive scenarios from each use case that specify refined operating conditions and the resulting intended behavior. Therefore, in this work, we propose to reference the graphs resulting from the behavior specification as the context, which specifies the operating conditions of the use

case and the maneuver under consideration. The graphical structure of the behavior specification has the advantage that deviations of the operating conditions can be investigated in a systematic and traceable way, hence supporting the hazard identification.

C. CAUSAL ANALYSIS (STPA)

The third step of the integrated approach is the causal analysis. The objective of this step is to identify the underlying causes of hazardous behavior. In the context of SOTIF, particular emphasis is placed on functional insufficiencies and triggering conditions. To identify these types of cause, STPA's focus on interactions and process/mental models is promising. However, as identified in Section III, STPA would benefit from SOTIF-tailored inputs. Specifically, a definition of operating conditions (Req. 3) and hazardous behavior (Req. 2) would help to apply STPA in the context of SOTIF. With the artifacts generated in the two previous steps, a SOTIF-tailored STPA can be executed. The behavior specification method provides relevant information on operating conditions, while the presented hazard identification enables linking STPA results to SOTIF objectives. To explain these interfaces in more detail, the SOTIF-tailored STPA will be briefly explained below.

The first step of STPA is the *definition of the analysis purpose*. When using STPA for SOTIF, the focus is on the loss of physical harm. Relevant hazards can be derived from the hazard identification approach. As the focus of ISO 21448 is on the hazardous behavior at the vehicle level, the preliminary system boundary should also be the vehicle.

In the second step of STPA, the *control structure is modeled*. The starting point for the control structure are the operating conditions (scene elements and inferred facts) from the behavior specification. At the same time, the control structure view adds relevant details for the causal analysis. STPA related additions include the hierarchic structuring and the addition of interactions.

After setting up the control structure, the third step of STPA is to *identify UCAs*. For this step, the information generated from the hazard identification process serves as a baseline. Each identified hazardous behavior can be translated into an UCA at vehicle level and should be analyzed in more detail.

To identify causes for identified UCAs, the fourth step of STPA *identifies loss scenarios* by inspecting the control structure. When identifying loss scenarios, we explicitly consider causal factors related to specification insufficiencies (e.g., missing consideration of covered lane markings), performance insufficiency (e.g., limited sensor resolution), and triggering conditions (e.g., snow or dirt on road). While not the primary focus of this manuscript, the identification of causal factors may also encompass aspects of misuse, which can be analyzed by inspecting interactions and mental models of humans included in the control structure as discussed in Section VII-B.

Compared to [11], we do not add a dedicated fifth step to identify triggering conditions and functional insufficiencies. Instead, the fourth STPA step is adjusted to explicitly characterize at least one triggering condition and at least one functional insufficiency for each loss scenario. As a result, each loss scenario matches the hazardous event model presented in Fig. 1. Hence, element-level causes can be traced to vehicle-level hazardous behavior. The level of detail of the identifiable causes depends on the granularity of the control structure. More guidance about the relation between vehicle-level behavior and element-level causes is provided in [40].

VI. EXAMPLE APPLICATION - PEOPLE MOVER BUS STOP

This section demonstrates the integrated approach with a simplified example. The objective of this section is to show the integration and not to argue scalability of the approach. It is important to note that the general approach does not necessarily require a specific model-based framework or profile and can be implemented with a custom solution. However, we utilize model-based design tools to implement the integrated approach and show how traceability can be established between STPA artifacts (e.g., causal factors) and SOTIF work products (especially triggering conditions and functional insufficiencies).

The example application requires the definition of an operational environment and an example system. As we intend to apply the integrated approach in a current research project, the example complies with the scope of the project. A concept vehicle analyzed in the project is *Ushift*, an automated people mover. The concept vehicle is developed by the German Aerospace Center (DLR) and is composed of a driveboard and a capsule, as explained by [43]. The driveboard provides the driving capabilities while the capsule is exchangeable to support either passenger transport or goods delivery. The selected use case focuses on the transport of passengers with a respective configuration of the vehicle. Ushift will drive in Braunschweig, Germany and will pick up and drop off people at regular bus stops in the city. In this article, we conduct an example analysis for a bus stop on the Tostmannplatz in Braunschweig, Germany. The bus stop is characterized by the layout and markings of the lane as depicted in Fig. 5.

A. SPECIFYING THE INTENDED BEHAVIOR

The use case comprises scenarios in which the U-Shift (our system of interest) approaches and departs from a bus stop. The behavior specification refines the operating conditions through the specification of maneuver options. We employ a graph-based representation of the behavior specification as outlined in Section V-A. The graph in Fig. 6 depicts the scene elements, their causal relationships, and the maneuvers intended when approaching the bus stop. When a pedestrian is waiting, the vehicle shall keep in the lane and stop at the bus stop. To determine whether a pedestrian is waiting, the vehicle must account for the validity of the bus stop, the position of pedestrians, and the relative distance of the pedestrian to the bus stop sign.

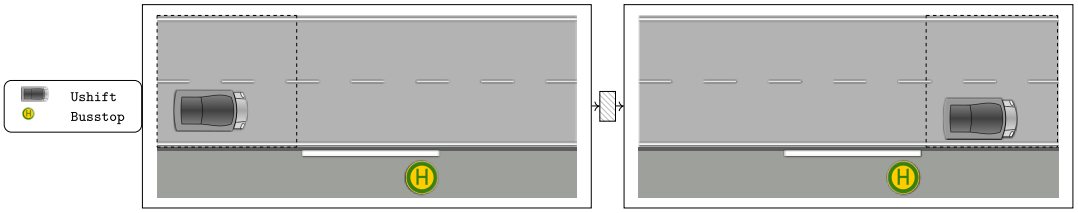


FIGURE 5. **Bus stop use case specified with traffic sequence charts [41], [42]**

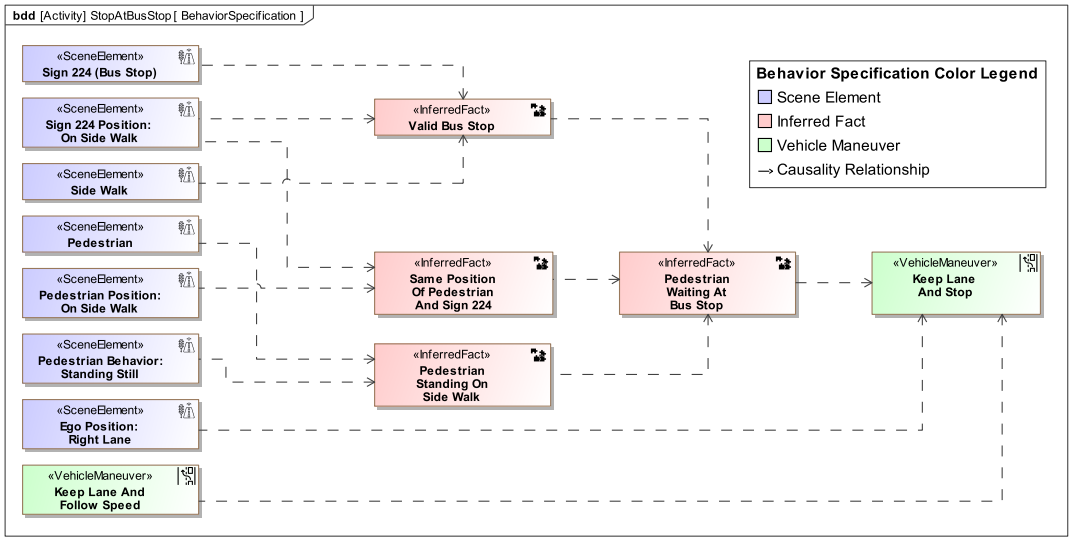


FIGURE 6. **Simplified specification of the intended behavior *StopAtBusStop***

B. IDENTIFYING HAZARDS AND HAZARDOUS BEHAVIOR

Based on the specification of the intended behavior, hazardous behaviors and hazards can be identified and analyzed using the approach described in Section V-B. The results are denoted in a tabular format as depicted in Fig. 7. Given the use case *Bus Stop*, the behavior specification *StopAtBusStop* presented in Fig. 6 refines the operating conditions under which the maneuver *Keep Lane And Stop* is intended. By applying keywords to the different nodes of the behavior specification, variations of the considered scenario are analyzed.

The first line of the table in Fig. 7 illustrates the investigation of a maneuver deviation. Applying the keyword *Provided incorrectly* to the given maneuver, we derive the (hazardous) behavior: *Vehicle does not keep the lane*. Given the operating conditions of the behavior specification *StopAtBusStop*, this behavior may lead the vehicle to drive over the curbstone and onto the sidewalk (observable effect). Since pedestrians are waiting at the bus stop (scenario condition), this can cause the vehicle to come too close to people (hazard) and result in a collision (hazardous event).

The second line of Fig. 7 presents a hazard derived from a deviation of the specified operating conditions considering the effects of the maneuver *Keep Lane And Stop*. Here, the keyword *Not Provided* is applied to the inferred fact *Valid Bus Stop* of the behavior specification *StopAtBusStop* that leads to the (hazardous) behavior: *Vehicle stops at invalid bus stop*. This behavior implies that the vehicle stops at an invalid location (observable effects). Such an unexpected stop may surprise the following traffic (scenario condition) and lead to a rear-end collision (hazardous event). In the same way, the procedure can be applied to every node in the behavior specification to systematically analyze the effects of deviating operating conditions.

C. IDENTIFYING CAUSES OF HAZARDOUS BEHAVIOR

The artifacts resulting from the two previous methods can be utilized as a starting point for a SOTIF-tailored causal analysis using STPA. The behavior specification defines operating conditions in terms of scene elements, intended behavior, and the rationale for the specified maneuver options (inferred facts). The hazard identification method provides a list of hazardous behaviors and hazards, which in turn requires a subsequent causal analysis. Moreover, each hazardous behavior is defined with specific operating conditions. Therefore, the context for STPA's causal analysis is explicitly defined. With the SOTIF-tailored input of the previous methods, the definition of the analysis purpose (first step of STPA) is covered. In ISO 21448, the loss to prevent is physical injury or damage to persons (harm). Consequently, related hazards were derived in Fig. 7 and are shown in Fig. 8. Finally, the

#	Name	Use Case	Maneu...	Correct In	Deviation of	Keyword	Hazardous Behavior	Observable Effects	Hazardous Event	Additional Scenario Conditions	Hazards
1	HI-1	Bus Stop	Keep Lane And Stop	StopAtBusStop	Keep Lane And Stop	Provided Incorrectly	Vehicle does not keep the lane.	Vehicle drives over curb stone and onto sidewalk.	Collision with pedestrians.	Pedestrians are close to lane.	Vehicle too Close to Infrastructure Vehicle too Close to People
2	HI-2	Bus Stop	Keep Lane And Stop	StopAtBusStop	Valid Bus Stop	Not Provided	Vehicle stops at invalid bus stop.	Vehicle holds at invalid location.	Collision with following traffic.	Other traffic is close behind the vehicle.	Vehicle too Close to Other Traffic

FIGURE 7. **Excerpt of hazard identification table**

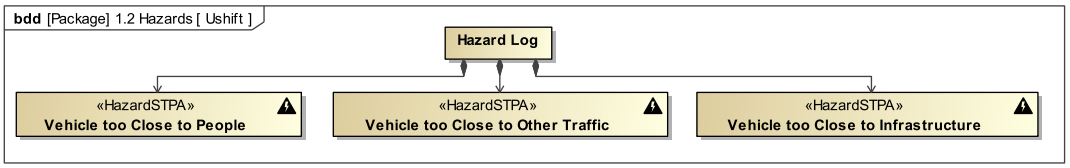


FIGURE 8. **Identified SOTIF related losses and hazards**

system boundary is specified at the vehicle level, as this is the level of interest for ISO 21448. To get more detail about the model-based STPA, interested readers are referred to previous work [44] and the corresponding GitHub repository[2].

In the second step of STPA, the control structure is created using the input of the behavior specification. For instance, scene elements involved in the behavior specification should be included in the control structure. Similarly, from the inference process, control actions, feedback, and process models can be derived and added. The control structure for the bus stop use case is shown in Fig. 9. The aspects of the behavior specification include elements such as *Passengers* and interactions such as *RoadSigns*.

In addition, other information can be added that is known about the vehicle and its operational environment from related systems engineering processes. This information is captured in the STPA's control structure as elements, interactions, and process/mental models. For example, a *Teleoperator* is added to Fig. 9 that assists the vehicle to operate using *OperationCMD*. Assumptions related to the *Teleoperator's* mental model are documented as properties such as the *Assumption about Supervised Vehicles*. Noting assumptions directly in the control structure model can help to identify corresponding loss scenarios. Assumptions should be documented and elaborated in more detail for reference during the analysis. One assumption is the *Assumption about Supervised Vehicles* which refers to the mental model of the *Teleoperator* about the amount of vehicles that he has to supervise. Another example related to the process model of the *Central Control System* is the *Assumption about Environment*. This assumption refers to the environmental operating conditions (weather, temperatures, etc.) that were considered within the system design. During the causal analysis, the influence and completeness of assumptions can be questioned in the context of unsafe interactions.

[2]https://github.com/DLR-FT/ModelBasedSTPA

Generally, the abstraction of the control structure can be adjusted depending on the focus of the analysis. In Fig. 9, the control structure not only focuses on the vehicle level, but also highlights an internal control loop. This view is chosen to show the possibility of connecting hazardous behavior at the vehicle level with the causal analysis of internal systems. After creating the control structure, the next step of STPA can be executed. As a starting point, the hazardous behaviors identified during the hazard identification method are translated into UCAs as shown in Fig. 10. The first UCA is derived from row one in Fig. 7 and describes the incorrect execution of *Maneuver*. The second UCA is derived from the second row of Fig. 7 and documents the stopping maneuver executed in a context where there is an invalid bus stop. By deriving both UCAs from the hazard identification, a traceability to the hazardous behavior and the resulting hazards is established.

To identify causes for hazardous behaviors and the corresponding UCAs, the control structure is analyzed in the fourth step of STPA. It is important to note that the focus is on SOTIF-related loss scenarios that explicitly define a triggering condition and functional insufficiency to comply with the hazardous event model in Fig. 1. Additional loss scenarios may be recognized [11], but these are beyond the scope of this work.

In this example, the analysis focuses on the *Maneuver* control action and the reasons why it is incorrectly provided during the bus stop scenario. Therefore, commands, feedback, and assumptions that contribute to the *Maneuver* are analyzed using the control structure. In addition, the corresponding behavior specification provides insight about relevant scene elements and inferred facts. Some of the derived loss scenarios are shown in Fig. 11. Each loss scenario description starts by documenting the hazardous behavior at vehicle level. In addition, the triggering condition and functional insufficiency, as well as their relationship, are mentioned. For example, in LS-1, the triggering condition is the *InfrastructurePerception*

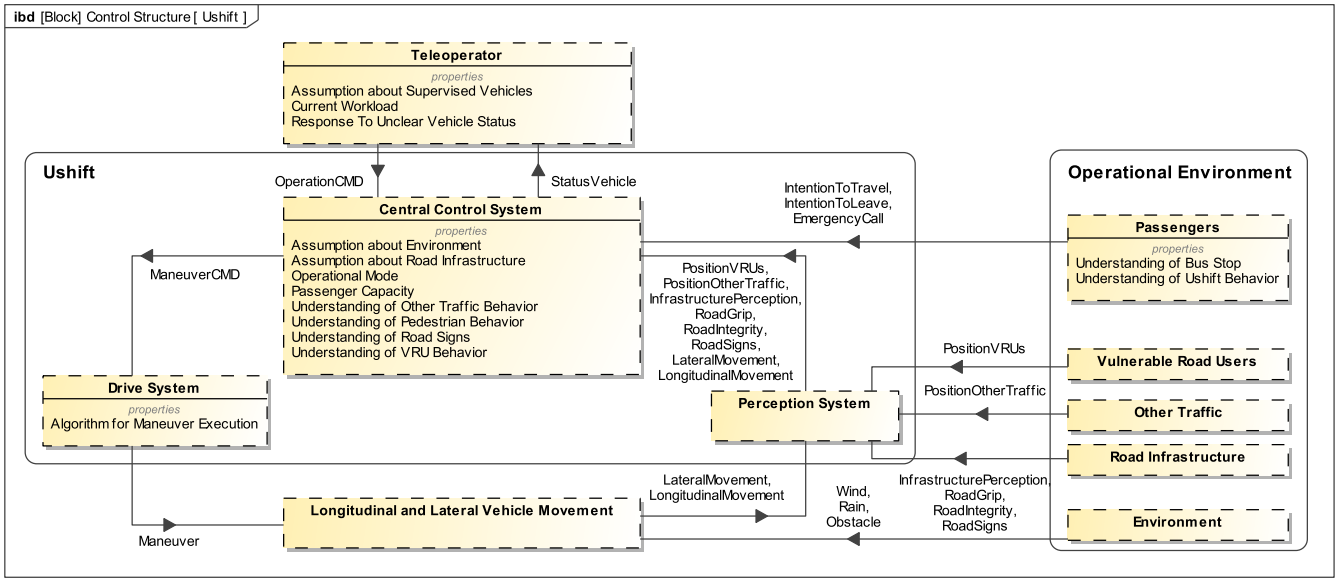


FIGURE 9. **Control structure with internal control loop of the Ushift**

#	Traced From	Name	UCA Description	Action	Hazard Reference	Loss Scenarios
1	HI-1	UCA-1	The [Vehicle] provides [Maneuver] to [Keep Lane and Stop Vehicle] incorrectly during the [StopAtBusStop] scenario when [Pedestrian Waiting At Bus Stop].	Maneuver	Vehicle too Close to Infrastructure Vehicle too Close to People	LS-1 LS-2
2	HI-2	UCA-2	The [Vehicle] provides the [Maneuver] to [Keep Lane and Stop Vehicle] during the [StopAtBusStop] scenario without a [Valid Bus Stop].	Maneuver	Vehicle too Close to Other Traffic	LS-3 LS-4

FIGURE 10. **UCAs derived from the hazard identification table**

feedback of a covered curb stone that is sent to *Central Control System*. The loss scenario manifests itself due to an inadequate *Assumption about Road Infrastructure* of the *Central Control System's* algorithm (functional insufficiency).

In summary, the example shows how an SOTIF-tailored STPA with a dedicated terminology mapping and methodological extensions can be used to systematically identify triggering conditions and functional insufficiencies. Fig. 12 describes some of the artifacts that are generated throughout the process as well as their relationship. Initially, we specified the intended behavior (e.g., *StopAtBusStop*) with the graph-based behavior specification. This behavior specification served as an input for the hazard identification method in which we analyzed maneuvers (e.g., *Keep Lane and Stop*) and operating conditions using a keyword-based approach (e.g., *Provided Incorrectly*). We translated deviations leading to hazards (e.g., *Vehicle too Close to Infrastructure*) into UCAs for the STPA-based causal analysis. Finally, we identified loss scenarios including triggering conditions (e.g., *Infrastructure Perception*) and functional insufficiencies (*Assumption about Road Infrastructure*).

VII. ASSESSMENT AND DISCUSSION

In this section, we evaluate the proposed methodology and application example with respect to the SOTIF objectives, discuss limitations of this article, and provide an outlook for future work.

A. MAPPING TO ISO 21448 OBJECTIVES

The proposed integrated approach has been specifically designed to address SOTIF. In particular, it has been designed to support the ISO 21448 objectives outlined in Section II-A.

The first step of the integrated approach – the behavior specification – contributes to Obj. 1. By defining use cases and specifying the intended behavior, as well as operating conditions, it provides information for the subsequent steps of the approach and thus for the following SOTIF activities. In addition, by using model-based design tools, the traceability of the different artifacts is ensured, which facilitates the update of the specification and design (Obj. 1).

The second step – the identification of hazards – is tailored to address Obj. 2 and 3. Hazardous behaviors are identified by systematic variations of the behavior specification. Further, by considering potential effects of hazardous behaviors,

#	Name	Loss Scenario Description	Triggering Condition	Functional Insufficiency
1	LS-1	The vehicle provides a wrong maneuver at the bus stop. The reason is that the central control system does not identify the lane markings correctly because the curb stone is covered (e.g. by dirt or now). Covered curb stones were not considered during the system design and lead to incorrect maneuver decisions.	InfrastructurePerception	Assumption about Road Infrastructure
2	LS-2	The vehicle provides a wrong maneuver at the bus stop. The reason is that the central control system calculates the maneuver command incorrectly on a day with harsh environmental conditions, because such conditions were not considered in the algorithm design.	Rain Wind	Assumption about Environment
3	LS-3	The vehicle stops at an invalid bus stop. The reason is that something closely resembles a bus stop sign and the central control system interprets the bus stop as valid.	RoadSigns	Understanding of Road Signs
4	LS-4	The vehicle stops at an invalid bus stop. The reason is that pedestrians are close on the sidewalk and the central control system interprets their behavior as intention to travel.	IntentionToTravel	Understanding of Pedestrian Behavior

FIGURE 11. **Loss scenarios with related triggering conditions and functional insufficiencies**

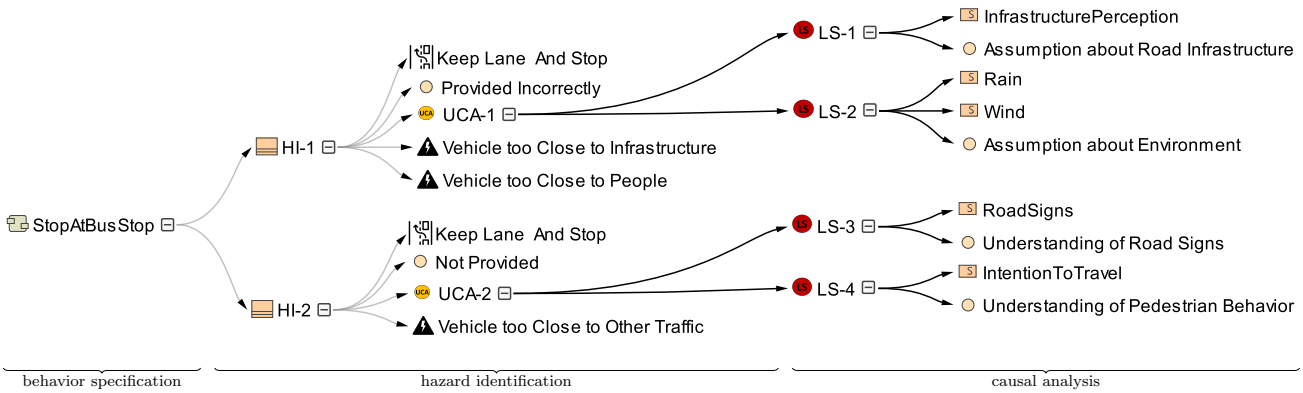


FIGURE 12. **Generated artifacts from behavior specification till causal analysis**

hazardous events and hazards are derived.

Finally, as proposed by ISO 21448, STPA is applied as a third step to identify functional insufficiencies and triggering conditions to achieve Obj. 4. The basis for this is given by mapping the different artifacts of STPA to SOTIF as discussed in Section IV. As STPA does not differentiate between the different causal factors, we have complemented the fourth step of STPA by adding a concrete designation of triggering conditions and functional insufficiencies when identifying loss scenarios as proposed by [11].

B. LIMITATIONS

ISO 21448 includes reasonably foreseeable direct and indirect misuse. Direct misuse is considered as a potential kind of triggering condition, while indirect misuse leads to reduced controllability of a hazardous event. In this article, we focused on triggering conditions that can be modeled as part of the operational domain. Since the operational domain covers humans, their interactions and process models, identification of direct misuse is supported by our methodology. Similarly, the identification of triggering conditions and functional insufficiencies that prevent the detection of indirect misuse can be realized. However, we did not focus on modeling and analyzing these two kinds of misuse in our example. To analyze related limitations, more extensive modeling of human-machine interaction as part of the behavior specification and control structure design is suggested.

Moreover, we used a simplified model of the operational domain for the behavior specification to illustrate the application of the integrated approach. To approach the identification of triggering conditions more systematically, more rigorous models of the operational domain should be applied (e.g., as proposed by [26], [45]). Similarly, we did not follow a specific methodology to generate the example behavior specification, as the focus of this article is the application of STPA given a behavior specification. In order to include stakeholder requirements, for example, from the traffic code systematically, [31] propose a method to derive behavior specifications.

Furthermore, ISO 21448 shows a strong focus on the identification of triggering conditions at the element level (e.g., conditions leading to radar reflections). Although STPA enables a hierarchical model of the control structure, we focused on the identification of triggering conditions at the vehicle level, as we identified terminological and methodological gaps at this level specifically. To identify element-level triggering conditions, a more detailed control structure perspective would be needed.

Finally, as with any scenario-based analysis, this article can only demonstrate scalability to a limited extent. All steps of the integrated approach involve assumptions, especially with respect to the complexity of the operational context and the

TABLE 5. ISO 21448 objectives mapped to steps of the integrated approach

ISO 21448 Objective	Supported By
Obj. 1 - System Specification and Design	Behavior Specification (step 1, Fig. 3)
Obj. 2 - Hazards	Hazard Identification Method (step 2, Fig 3)
Obj. 3 - Hazardous Behavior	Hazard Identification Method (step 2, Fig 3)
Obj. 4 - Functional insufficiencies and triggering conditions	Causal Analysis (step 3, Fig 3)

complexity of the system considered.

C. FUTURE WORK

Based on our findings and the highlighted limitations, future work could, investigate how reasonably foreseeable misuse can be integrated as part of the integrated approach. In general, STPA often considers humans as part of the control structure [5], [20], indicating that STPA could be used for this purpose. To improve the applicability of the STPA-based causal analysis for misuse identification, a dedicated representation and assessment of the mental models of involved humans is beneficial. Guidance for extending STPA for human analysis is, e.g., shown in [46].

Second, a systematic formalization of a domain metamodel for both the operational domain and, e.g., the system architecture would help to increase consistency for the implementation of the integrated approach [19], [44]. The generation of such (semi-)formal models would also allow partial automation of the integrated approach.

Finally, since the application scenario in Section VI was simplified and mainly served the purpose of demonstrating the method, the scalability of the approach should be further investigated in future work. For instance, the completeness of behavior specification and hazard identification cannot be guaranteed for an automated vehicle in an open context [47], [48]. Therefore, a more detailed application of the integrated approach for a defined operational domain, including an assessment of the residual risk, can support the effectiveness evaluation.

VIII. CONCLUSION

In this article, we tackle the terminological and methodological gaps that occur when STPA is applied in the context of ISO 21448. One of the main challenges for a consistent application of ISO 21448 is that some SOTIF related concepts, such as hazardous behavior, leave room for interpretation. To tackle this challenge, we provided a terminology mapping from STPA concepts to SOTIF concepts. For example, we showed that hazardous behavior (from SOTIF) can only be mapped to STPA's UCA, if the system boundary of the UCA is defined at the vehicle level. Furthermore, our analysis confirms [11] with their finding that the causal factors of STPA can be triggering conditions as well as functional insufficiencies, and need to be specified separately to create the necessary SOTIF work products. To implement a SOTIF-tailored STPA, we propose to explicitly include both triggering conditions and functional insufficiencies in loss scenarios.

Based on this terminology mapping, we propose methodological extensions that combine STPA with two additional approaches. First, we include a semiformal specification of the operational environment and vehicle behavior to provide a baseline for identifying hazardous behavior. Second, we propose a systematic hazard identification based on the specification of the operational environment to provide a SOTIF-focused input for the STPA application. As a result, we showed how the combination of these methods can be applied to create STPA results that can be mapped to relevant ISO 21448 objectives and prescribed SOTIF terminology. To do this, we applied the integrated approach to define and analyze a bus stop scenario of a concept vehicle. All steps of the integrated approach were executed in a model-based environment to increase traceability of the results.

Throughout the application of the integrated approach we made additional observations. First, STPA can benefit from additional guidance in defining domain-specific control structures [26]. Otherwise, the quality of the safety analysis is fundamentally dependent on expert knowledge. In this article, we showed that both control actions and the considered elements can be supported by the systematic integration of a formalized description of the operational environment. Second, given a SOTIF tailored control structure, the causal analysis provided by STPA supports a systematic identification of triggering conditions and functional insufficiencies as requried by ISO 21448. Finally, a model-based execution of the behavior specification and safety analysis can improve the traceability of the resulting artifacts.

References

[1] M. Martínez-Díaz and F. Soriguera, "Autonomous vehicles: Theoretical and practical challenges," *Transportation Research Procedia*, vol. 33, pp. 275–282, 2018, ISSN: 2352-1465. DOI: 10.1016/j.trpro.2018.10.103.

[2] N. G. Leveson and J. P. Thomas, "Certification of safety-critical systems," *Communications of the ACM*, vol. 66, no. 10, pp. 22–26, Sep. 2023, ISSN: 0001-0782. DOI: 10.1145/3615860.

[3] N. Annable, A. Bayzat, Z. Diskin, M. Lawford, R. Paige, and A. Wassyng, "Model-driven safety of autonomous vehicles," in *Recent Trends and Advances in Model Based Systems Engineering*, A. M. Madni, B. Boehm, D. Erwin, M. Moghaddam, M. Sievers, and M. Wheaton, Eds., Cham: Springer International Publishing, 2022, pp. 407–417, ISBN: 978-3-030-82083-1. DOI: 10.1007/978-3-030-82083-1_34.

[4] ISO 21448:2022, *Road vehicles – Safety of the intended functionality*, Standard, 2022.

[5] N. G. Leveson and J. P. Thomas, "STPA Handbook," 2018. Accessed: Oct. 14, 2025. [Online]. Available: http://psas.

scripts.mit.edu/home/get_file.php?name=STPA_Handbook.pdf.

[6] E. Acar Celik, C. Cârlan, A. Abdulkhaleq, F. Bauer, M. Schels, and H. J. Putzer, "Application of STPA for the elicitation of safety requirements for a machine learning-based perception component in automotive," in *Computer Safety, Reliability, and Security*, M. Trapp, F. Saglietti, M. Spisländer, and F. Bitsch, Eds., Cham: Springer International Publishing, 2022, pp. 319–332, ISBN: 978-3-031-14835-4. DOI: 10.1007/978-3-031-14835-4_21.

[7] S. Khastgir, S. Brewerton, J. Thomas, and P. Jennings, "Systems approach to creating test scenarios for automated driving systems," *Reliability Engineering & System Safety*, vol. 215, p. 107 610, 2021, ISSN: 0951-8320. DOI: 10.1016/j.ress.2021.107610.

[8] I. Friedberg, K. McLaughlin, P. Smith, D. Laverty, and S. Sezer, "STPA-SafeSec: Safety and security analysis for cyber-physical systems," *Journal of Information Security and Applications*, vol. 34, pp. 183–196, Jun. 2017, ISSN: 2214-2126. DOI: 10.1016/j.jisa.2016.05.008. Accessed: Oct. 14, 2025.

[9] ISO/SAE 21434:2021, *Road vehicles – Safety of the intended functionality*, Standard, 2021.

[10] ISO/TS 5083, *Road vehicles — Safety for automated driving systems — Design, verification and validation*, Technical Specification, 2025.

[11] R. Graubohm, M. Loba, M. Nolte, and M. Maurer, "Identifikation Auslösender Umstände von SOTIF-Gefährdungen durch Systemtheoretische Prozessanalyse (in German)," *at - Automatisierungstechnik*, vol. 71, no. 3, pp. 209–218, 2023. DOI: doi:10.1515/auto-2022-0164. Accessed: Oct. 14, 2025.

[12] R. Wang and R. Niu, "Research on the SOTIF analysis method for autonomous perception function based on adapted STPA," in *Proceedings of the 6th International Conference on Electrical Engineering and Information Technologies for Rail Transportation (EITRT) 2023*, J. Yang, D. Yao, L. Jia, Y. Qin, Z. Liu, and L. Diao, Eds., Singapore: Springer Nature Singapore, 2024, pp. 535–548, ISBN: 978-981-9993-15-4. DOI: 10.1007/978-981-99-9315-4_52.

[13] L. Putze, L. Westhofen, T. Koopmann, E. Böde, and C. Neurohr, "On quantification for SOTIF validation of automated driving systems," in *2023 IEEE Intelligent Vehicles Symposium (IV)*, 2023, pp. 1–8. DOI: 10.1109/IV55152.2023.10186795.

[14] ISO 26262:2018, *Road vehicles – Functional safety*, Standard, 2018.

[15] ISO/IEC Guide 51: 2014, *Safety aspects — Guidelines for their inclusion in standards*, Standard, 2014.

[16] S. Ulbrich, T. Menzel, A. Reschka, F. Schuldt, and M. Maurer, "Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, Gran Canaria, Spain, Sep. 2015, pp. 982–988, ISBN: 978-1-4673-6596-3. DOI: 10.1109/ITSC.2015.164.

[17] SAE, *J3307_202503: System theoretic process analysis (STPA) standard for all industries*, 2025. DOI: https://doi.org/10.4271/J3307_202503.

[18] SAE J3016:2021, *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*, Standard, 2021.

[19] M. Nolte and M. Maurer, "Towards Closing the Gap between Model-Based Systems Engineering and Automated Vehicle Assurance: Tailoring Generic Methods by Integrating Domain Knowledge," in *16. Workshop Fahrerassistenz und Automatisiertes Fahren*, Irsee, Germany: Uni-DAS e. V., 2025. [Online]. Available: https://www.uni-das.de/images/pdf/fas-workshop/2025/FAS2025-12-Nolte-Maurer.pdf.

[20] N. G. Leveson, *Engineering a Safer World. Systems Thinking Applied to Safety* (Engineering Systems). Cambridge: The MIT Press, 2016.

[21] L. Junfeng, Z. Yunshuang, Z. Shuai, C. Chao, and D. Zhibin, "A research on SOTIF of LKA based on STPA," in *2022 IEEE International Conference on Real-Time Computing and Robotics (RCAR)*, 2022, pp. 396–400. DOI: 10.1109/RCAR54675.2022.9872242.

[22] S. Zhang, T. Tang, and J. Liu, "A hazard analysis approach for the SOTIF in intelligent railway driving assistance systems using STPA and complex network," *Applied Sciences*, vol. 11, no. 7714, 2021, ISSN: 2076-3417. DOI: 10.3390/app11167714.

[23] SAE, *J3187_202305: System theoretic process analysis (STPA) recommended practices for evaluations of safety-critical systems in any industry*, 2023. DOI: 10.4271/j3187_202305.

[24] J. Thomas, "Extending and automating a systems-theoretic hazard analysis for requirements generation and analysis," Ph.D. dissertation, Massachusetts Institute of Technology, 2023. Accessed: Oct. 14, 2025. [Online]. Available: https://dspace.mit.edu/bitstream/handle/1721.1/81055/857791969-MIT.pdf.

[25] M. Oldoni and S. Khastgir, "Introducing ODD-SAF: An operational design domain safety assurance framework for automated driving systems," in *Road Vehicle Automation 10*, G. Meyer and S. Beiker, Eds., Cham: Springer Nature Switzerland, 2023, pp. 133–151, ISBN: 978-3-031-34757-3. DOI: 10.1007/978-3-031-34757-3_11.

[26] T. Nakashima, R. Kureta, and S. Khastgir, "Addressing systemic risks in autonomous maritime navigation: A structured STPA and ODD-based methodology," *Reliability Engineering & System Safety*, vol. 261, p. 111 041, 2025, ISSN: 0951-8320. DOI: https://doi.org/10.1016/j.ress.2025.111041.

[27] G. Bagschik, T. Menzel, and M. Maurer, "Ontology based scene creation for the development of automated vehicles," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018, pp. 1813–1820. DOI: 10.1109/IVS.2018.8500632.

[28] M. Scholtes et al., "6-layer model for a structured description and categorization of urban traffic and environment," *IEEE Access*, vol. 9, pp. 59 131–59 147, 2021. DOI: 10.1109/ACCESS.2021.3072739.

[29] P. Koopman and W. Widen, "Safety ethics for design and test of automated driving features," *IEEE Design & Test*, vol. 41, no. 1, pp. 17–24, 2024. DOI: 10.1109/MDAT.2023.3281733.

[30] N. F. Salem et al., "Safety and Risk – Why their Definitions Matter," in *Handbook Assisted and Automated Driving*, ser. ATZ/MTZ-Fachbuch, H. Winner, K. Dietmayer, L. Eckstein, M. Jipp, M. Maurer, and C. Stiller, Eds., to be published., Wiesbaden: Springer Vieweg.

[31] N. F. Salem et al., "An ontology-based approach toward traceable behavior specifications in automated driving," *IEEE Access*, vol. 12, pp. 165 203–165 226, 2024. DOI: 10.1109/ACCESS.2024.3494036.

[32] I. Jatzkowski et al., "Zum Fahrmanöverbegriff im Kontext Automatisierter Straßenfahrzeuge (in German)," TU Braunschweig, Braunschweig, Technischer Bericht, 2021. [Online]. Available: https://www.tu-braunschweig.de/fileadmin/Redaktionsgruppen/Institute_Fakultaet_5/IFR/Dateien_EFS/Jatzkowski_et_al._-_Zum_Fahrmanoeverbegriff_im_Kontext_automatisierter_.pdf.

[33] N. F. Salem et al., "Risk Management Core—Toward an Explicit Representation of Risk in Automated Driving," *IEEE Access*, vol. 12, pp. 33 200–33 217, 2024, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2024.3372860.

[34] M. Scholtes et al., "6-Layer Model for a Structured Description and Categorization of Urban Traffic and Environment,"

IEEE access : practical innovations, open solutions, vol. 9, pp. 59 131–59 147, 2021. DOI: 10.1109/ACCESS.2021.3072739.

[35] K. Czarnecki, "Operational World Model Ontology for Automated Driving Systems - Part 2: Road Users, Animals, Other Obstacles, and Environmental Conditions," Waterloo Intelligent Systems Engineering (WISE) Lab, University of Waterloo, Waterloo, Canada, Technical Report, Jul. 2018.

[36] K. Czarnecki, "Operational World Model Ontology for Automated Driving Systems - Part 1: Road Structure," Waterloo Intelligent Systems Engineering (WISE) Lab, University of Waterloo, Waterloo, Canada, Technical Report, Jul. 2018. DOI: 10.13140/RG.2.2.15521.30568.

[37] P. Irvine, X. Zhang, S. Khastgir, and P. Jennings, "Structured Natural Language for expressing Rules of the Road for Automated Driving Systems," in *2023 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2023.

[38] BSI Flex 1891 v1.0, "Behaviour taxonomy for automated driving system (ADS) applications – Specification," Flexible Standard, 2025.

[39] B. Kramer, C. Neurohr, M. Büker, E. Böde, M. Fränzle, and W. Damm, "Identification and quantification of hazardous scenarios for automated driving," in *Model-Based Safety and Assessment*, M. Zeller and K. Höfig, Eds., Cham: Springer International Publishing, 2020, pp. 163–178. DOI: 10.1007/978-3-030-58920-2_11.

[40] K. Czarnecki, H. Kuwajima, K. Czarnecki, and H. Kuwajima, "STEAM & MoSAFE: SOTIF Error-and-Failure Model & Analysis for AI-Enabled Driving Automation," in *WCX SAE World Congress Experience*, SAE International, 2024. DOI: 10.4271/2024-01-2643. Accessed: Oct. 14, 2025.

[41] W. Damm, S. Kemper, E. Möhlmann, T. Peikenkamp, and A. Rakow, "Using Traffic Sequence Charts for the Development of HAVs," in *ERTS 2018*, ser. 9th European Congress on Embedded Real Time Software and Systems (ERTS 2018), Toulouse, France, Jan. 2018. [Online]. Available: https://hal.science/hal-01714060.

[42] W. Damm, E. Möhlmann, T. Peikenkamp, and A. Rakow, "A Formal Semantics for Traffic Sequence Charts," in *Principles of Modeling: Essays Dedicated to Edward A. Lee on the Occasion of His 60th Birthday*, M. Lohstroh, P. Derler, and M. Sirjani, Eds. Cham: Springer International Publishing, 2018, pp. 182–205, ISBN: 978-3-319-95246-8. DOI: 10.1007/978-3-319-95246-8_11.

[43] M. Münster et al., "U-Shift II vision and project goals," in *22. Internationales Stuttgarter Symposium*, M. Bargende, H.-C. Reuss, and A. Wagner, Eds., Wiesbaden: Springer Fachmedien Wiesbaden, 2022, pp. 18–31, ISBN: 978-3-658-37011-4. DOI: 10.1007/978-3-658-37011-4_3.

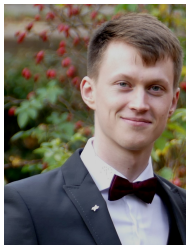
[44] A. Ahlbrecht, J. Sprockhoff, and U. Durak, "A System-Theoretic Assurance Framework for Safety-Driven Systems Engineering," *Software and Systems Modeling*, vol. 42, pp. 253–270, 2024, ISSN: 2169-3536. DOI: 10.1007/s10270-024-01209-6.

[45] L. Westhofen, C. Neurohr, M. Butz, M. Scholtes, and M. Schuldes, "Using Ontologies for the Formalization and Recognition of Criticality for Automated Driving," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 519–538, 2022, ISSN: 2687-7813. DOI: 10.1109/OJITS.2022.3187247.

[46] M. E. France, *Engineering for humans: A new extension to STPA*, 2017. Accessed: Oct. 14, 2025. [Online]. Available: https://dspace.mit.edu/handle/1721.1/112357.

[47] J. E. Stellet, T. Brade, A. Poddey, S. Jesenski, and W. Branz, "Formalisation and algorithmic approach to the automated driving validation problem," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2019, pp. 45–51. DOI: 10.1109/IVS.2019.8813894.

[48] S. Burton, I. Habli, T. Lawton, J. McDermid, P. Morgan, and Z. Porter, "Mind the gaps: Assuring the safety of autonomous systems from an engineering, ethical, and legal perspective," *Artificial Intelligence*, vol. 279, p. 103 201, Feb. 2020, ISSN: 0004-3702. DOI: 10.1016/j.artint.2019.103201.



ALEXANDER AHLBRECHT works as a scientific associate in the Institute of Flight Systems at the German Aerospace Center (DLR e.V.). He received his master's degree in Electronic Automotive and Aerospace Systems at the TU-Braunschweig and is a PhD candidate at the TU-Clausthal. His research focuses on the integration of Model-Based Systems Engineering (MBSE) and Model-Based Safety Assessment (MBSA) for the development of safety-critical systems.

NAYEL FABIAN SALEM received the B.Sc.degree in mechanical engineering from Technische Universität Berlin, Berlin, Germany, in 2018, and the M.Sc. degree in electronic systems engineering from the Technische Universität Braunschweig, Braunschweig, Germany, in 2020, where he is currently pursuing the Ph.D. degree with the Institute of Control Engineering. Since 2021, he has been a Research Associate. His main research interests include safety assurance of automated vehicles, focusing on safety issues regarding behavior specification.

LINA PUTZE received the B.Sc. and M.Sc. degrees in mathematics from the University of Münster in 2016 and 2019, specializing on the topics of stochastic processes, probability theory and its applications. She is currently working as a researcher at the group System Concepts and Design Methods at the German Aerospace Center (DLR e.V.) Institute of Systems Engineering for Future Mobility. The focus of her research is on methods to ensure trustworthiness of highly automated transport systems in different domains, including the identification and analysis of hazards and risk triggering scenario properties, causal analysis and risk assessment.

INGO STIERAND is a research scientist at the German Aerospace Center (DLR e.V.), Institute of Systems Engineering for Future Mobility in Oldenburg. His work focuses on model-based systems engineering (MBSE), safety and reliability analysis, and the design of dependable real-time systems. He is particularly interested in contract-based design approaches and their application to complex cyber-physical and autonomous systems. His contributions cover several national and international research projects and regularly publishes in the fields of systems engineering, formal methods, and automotive safety.

UMUT DURAK is the Group Leader for Avionics Systems in the Institute of Flight Systems at the German Aerospace Center (DLR). He is also a Professor for Aeronautical Informatics in the Informatics Institute at the TU-Clausthal. His research interests concentrate on engineering of software intensive airborne systems. He has published 5 books and more than 80 papers in various conference proceedings and journals. He is an Associate Fellow and the Co-Chair of Software Technical Committee at the American Institute of Aeronautics and Astronautics (AIAA) and an Executive Board Member of the German simulation association Arbeitsgemeinschaft Simulation (ASIM).

MARCUS NOLTE is a Post-Doc in the Department of Engineering Design, Unit Mechatronics, at KTH Royal Institute of Technology, Stockholm. Before, he was a PostDoc at the Institute of Control Engineering at TU Braunschweig, Germany. There he received his PhD in Systems and Safety Engineering for Automated Vehicles. He holds a M.Sc. in Electrical Engineering with a focus on Mechatronics and Control from TU Braunschweig. His current research focuses on safety and trustworthiness assurance for AI-based systems with a particular emphasis on automated vehicles and the domain-specific tailoring of Model-Based Systems and Safety Engineering methods.

ECKARD BÖDE received his Dipl.-Inform. degree in Computer Science from the Carl von Ossietzky University, Oldenburg, Germany, in 2001. He subsequently joined OFFIS e.V., where he focused on safety assessment and model-based safety analysis for aerospace and automotive applications. In 2012, he was appointed Group Leader for Safety Analysis and Verification. He currently leads the R&D group System Concepts and Design Methods at the German Aerospace Center (DLR e.V.), Institute of Systems Engineering for Future Mobility. His research interests include methods and tools for the design and verification of trustworthy cyber-physical systems, with a particular emphasis on safety assessment of automated systems and the integration of functional safety with SOTIF in safety cases.