Driver assistance systems: analysis, tests and the safety case. ISO 26262 and ISO PAS 21448

O M Kirovskii and V A Gorelov
Department of Tracked Vehicles and Mobile Robots, Bauman Moscow State Technical University, 2nd Baumanskaya st. 5, 105005 Moscow, Russian Federation

Abstract. Modern automotive system must comply with strict safety requirements. This paper focuses on two aspects of safety: functional safety per ISO 26262 (FS) and the safety of the intended functionality (SOTIF) per ISO PAS 21448. The FS encompasses a lifecycle ensuring the absence of unreasonable risks due to internal failures of the system. SOTIF concentrates on non-deterministic parts and algorithms (e.g. neural networks), as full specification of their performance is out of reach now. At the same time, FS and its lifecycle are better known to the community and have better history of implementation. This paper presents an approach to integration of FS and SOTIF requirements based on FS lifecycle.

1. Introduction
Increasing autonomy of road vehicles is one of the most important trends shaping today’s automotive industry. One of the challenges that developers and OEMs face is the requirement to guarantee safety and security to people in and around the autonomous vehicle (AV) under all circumstances. The methods for safety assurance are constantly evolving for many years already, without any final stop in sight. This introduction gives a brief definition of the terms related to automotive safety and main strategies the engineers pursue in assuring it.

According to the Technical Regulation of the Eurasian Economic Union TR TS 018/2011 “Safety of Wheeled Vehicles” [1], “safety of the vehicle” is defined as “state (…) where the risk of harm to life, health of people as well as to the property (…) is eliminated or minimized”. The lawmakers acknowledge the existence of conditions under which complete elimination of risks related to the vehicle is impossible. The intent of the law is to motivate the vehicle designers and manufacturers and other vehicle lifecycle participants to take measures against those risks through the entire lifecycle.

Norms on functional safety (e.g. ISO 26262 [2]) share this paradigm. Functional safety (FS) addresses faults resulting from malfunctions of systems (i.e. failures of its components). Modern autonomous driving (AD) systems and advanced driver assistance systems (ADAS) highly depend on the data from external sources provided by non-deterministic, often AI-based algorithms (e.g. object recognition, path planning using machine learning techniques). In those cases, even a fault-free system may exhibit hazardous behavior due to the limitations of the system’s performance. The new paradigm of Safety of the Intended Functionality (SOTIF) as formulated in ISO PAS 21448 [3] addresses the issue.

Both FS and SOTIF prescribe a lifecycle process for identification and control of risks. The FS lifecycle is better known and more widely implemented because of its longer history, while SOTIF measures are becoming necessary with increasing autonomy of vehicles. The two lifecycles can be
combined in order to achieve throughout traceability of requirements and facilitate creation of the Safety Case.

2. Scope of FS and SOTIF
It is known that one may find more than 40 definitions of safety only among different ISO standards. Surprisingly, both ISO 26262 and ISO PAS 21448 agree on this issue and define safety as absence of unreasonable risks. Based on the definition, the scope of each standard is defined:

- Functional safety: absence of unreasonable risk due to hazards caused by malfunctioning behavior of E/E Systems (ISO 26262-1, Def. 1.51)
  - Malfunctioning behavior is defined as a failure or unintended behavior of an item with respect to its design intent (which in turn is undefined).
- Safety of the Intended Functionality: Absence of unreasonable risk due to hazards caused by performance limitations of the intended behavior or by reasonable foreseeable misuse by the user (ISO PAS 21448, Def. 3.5)
  - Performance limitation: insufficiencies of the function itself (example: incorrect perception of the situation or caused by model assumptions that are not feasible) – ISO PAS 21448, Def. 3.4

The scopes defined in both standards are not separated clear enough, as there is no clear distinction provided between “malfunction” and “performance limitation”. Both are defined as deviations from desired behavior. It can be said that there is an overlap between the scope of the FS and SOTIF. Instead of refining the definitions in order to draw a clearer distinction, a more useful approach would be to define the aspects of safety which are best treated with FS methodology and those where SOTIF methodology is more beneficial.

![Safety Methodology: Use Cases](image-url)

**Figure 1.** Safety methodology: use cases.
Safety methodologies are employed to control (i.e. systematically identify, reduce, mitigate and assess the residuals) risks. Different methodologies here strive for the same goal. Meaningful use of various methodologies together may be achieved only if methods related to different methodologies are put together into one lifecycle.

Figure 1 comes without a claim for completeness. Besides that, cross-dependencies exist, e.g. functional safety addresses unintentional misuse if a misuse case becomes hazardous in case of malfunctions; cybersecurity also handles the misuse cases which are not entirely intentional, e.g. in the area of human-machine interface (HMI). This paper will address FS and SOTIF lifecycle; cybersecurity is not addressed. However, the approach defined here is capable of adapting cybersecurity and possibly other safety methodologies (i.e. safety of mechanical components, chemical safety etc.)

3. Generalized Safety Lifecycle (FS + SOTIF)

3.1. ISO 26262

The lifecycle of functional safety is presented on the well-known “3-V diagram” in the ISO 26262.

![Safety lifecycle according to ISO 26262:2011](image)

The numbers 1 to 10 mark the parts of the standard relevant to each lifecycle item. The lifecycle consists of five phases: concept phase (Part 3 of the standard), product development on the system (Part 4), HW (Part 5) and SW (Part 6) levels as well as production and operation (Part 7). For now, SOTIF is limited to the concept phase and system-level activities. Technical details of ensuring the safety of the intended functionality on SW level are part of what may be called “Safety of Machine Learning (ML)”. Currently those aspects are not addressed within the ISO PAS 21448.
3.2. ISO PAS 21448

SOTIF lifecycle consists of three major steps:

1. “Evaluate by analysis”: identification and evaluation of **triggering events** (i.e. specific conditions of a driving scenario that serve as an initiator for a subsequent system reaction possibly leading to a hazardous event, see [3], def. 3.15).

2. “Evaluate known hazardous scenarios”: identification of relevant hazardous scenarios from the set of known hazardous scenarios, explicit hazard testing

3. “Evaluate unknown hazardous scenarios”: “statistical testing”, i.e. search validation via exposure of the system to the statistically distributed conditions of real world or simulation.

ISO PAS 21448 consists of 12 chapters which run along the V-lifecycle close to the lifecycle of the ISO 26262. Fig. 10 in [3] provides a link from SOTIF lifecycle to FS lifecycle. However, this link is too general and therefore not useful. Table 1 provides a view of a lifecycle integrating FS and SOTIF requirements (generalized safety lifecycle).

4. Generalized Safety Lifecycle (a suggestion)

In Table 1, *italicized* items show work products (WP) existing in FS lifecycle and getting additional requirements from SOTIF lifecycle. Triggering Event Analysis (TEA), the WP highlighted *bold*, is the only SOTIF-specific WP.

Of 14 WPs in Generalized Safety Lifecycle, one (TEA) is newly introduced due to SOTIF and 3 do not get any additional requirements. The residual 10 WPs contain both FS and SOTIF requirements.

**Table 1.** Generalized safety lifecycle: FS + SOTIF.

| Chapter | Work product | Chapter | Work product | Generalized Safety Lifecycle (work products) |
|---------|--------------|---------|--------------|---------------------------------------------|
| 3-5     | Item Definition | 5       | Functional and System Specification | Item Definition |
| 3-7     | Hazard Analysis and Risk Assessment | 6       | Hazard Identification, Hazard Analysis, Risk Evaluation, Specification of Validation Target | Hazard Analysis and Risk Assessment |
|         |               | 7       | Identification of Triggering Events, Assessment of the Acceptability of Triggering Events | Triggering Event Analysis (TEA) |
| 3-8     | Functional Safety Concept |         |               |                                             |
| 9       | Definition of V&V Strategy |         | SOTIF V&V Strategy |


4-6.4.6 System Verification Report 10 SOTIF Verification Report System Verification Report
4-6.4.6.2 Validation Plan
4-7.4 Technical Safety Concept Technical Safety Concept
4-7.4 System Design Specification System Design Specification
4-7.4.3 Safety Analysis Report Safety Analysis Report
4-8.4 Integration Testing Plan & Report Integration Testing Plan and Report
4-9.4 Validation Plan & Report 11 SOTIF Validation Report Validation Plan & Report
4-10.4 Functional Safety Assessment Report
4-11.4 Release for Production Report 12 SOTIF Release Report Safety Analysis Report

The next chapter of this paper goes through specific work products of the Generalized Safety Lifecycle and presents the integration of the FS and SOTIF lifecycles on the level of work products.

5. Integration of the Lifecycles
Now we have defined the work products which should be changed or where additional information needs to be added in order to integrate the requirements of ISO PAS 21448. The next question is: which information needs to be added to each of those WPs and to the safety lifecycle as a whole? What are the requirements of the SOTIF lifecycle?

The paper [4] presents the general safety case (SC) for ML. Even though SOTIF is not equal to “automotive ML applications”, it is possible to use the safety case model of the paper to cover SOTIF-related safety cases as well.

The goal of the ML safety case (according to [4]): to prove that the residual risk associated with functional insufficiencies in the object detection and classification function is acceptable. The argument here shall be provided over the space of assumptions on the operational profile of the system, inputs to the ML functions and performance limitations of its algorithms.

The assumptions contain data which are not project-specific and therefore not part of the project’s safety case. They may be fully or partially shared between projects. Each project may use the whole set of available data or a sub-set. The data may be found via safety analyses which are part of standard safety lifecycle as well as via SOTIF-specific steps of the safety lifecycle. The only SOTIF-specific WP identified in Table 1 is the Triggering Event Analysis (TEA). However, another allocation of SOTIF requirements may result in other WPs characterizing the SOTIF-specific part of the lifecycle.
The Generalized Safety Lifecycle as defined in §4 is represented graphically on the Figure 3. It consists of functional safety lifecycle according to ISO 26262 with some WPs receiving requirements from ISO PAS 21448 alongside with normal safety requirements. The safety lifecycle is extended via SOTIF lifecycle, currently containing only the procedure of the Triggering Event Analysis. Both FS and SOTIF parts of the Generalized Safety Lifecycle are based on the data on the intended functionality.

The data on the intended functionality include the information currently known about the functionality to be implemented in the project. The data are project-independent; therefore, it may be stored centrally in a databank within an organization or an organizational division.

The necessary data on the intended functionality include:
- Triggering events
- Known relevant use cases, both safe and hazardous, with references to the relevant triggering events
- Statistics on the occurrence of the known relevant use cases
- Expected human reactions in hazardous cases. This is specifically needed for the systems featuring SAE automation levels of 1 to 3, as the safety case for those systems typically includes human driver as a safety measure.

The databank may include more data, depending on the data model implemented within the organization.

5.1. Hazard Analysis and Risk Assessment

The HARA consists of hazard identification and hazard evaluation.

Hazard identification is an inductive process aimed at the identification of possible hazards as well as risk related to them. Hazard identification always takes into account the situation (use case, driver’s and road conditions etc.), which may be considered external to the system and the deviation of the actual system behavior from some expected behavior. Generalized Safety Lifecycle follows the FS lifecycle in consideration of functional failures (malfunctions) as deviations for hazard identification.

Hazard evaluation happens differently within FS and SOTIF frameworks. This is caused by the different understanding of the “reasonable” or target levels of risk in the two standards.

ISO 26262 features four Automotive Safety Integrity Levels (ASILs) as well as “QM” safety level for the cases where the residual risk is considered acceptable to be minimized by quality management approaches without implementation of any specific safety measures. ASIL levels play a dual role: they measure risks associated with a specific malfunction in a given situation (with ASIL A for the lowest
risk and ASIL D for the highest), and at the same time ASIL allocation defines the requirements of the standard applicable to the system, i.e. the amount of effort to be invested in order to minimize the existing risk to a tolerable level. Thus, both possible levels of risk to be found in the system as well as the tolerable level are given by the standard itself, even though this information is implicit (unlike e.g. IEC 61508 where quantitative risk assessment is possible).

According to ISO 26262, only the most hazardous use case is relevant for the definition of the maximum ASIL for the relevant safety goal. Then, the defined safety goal is used as a proxy for all technical activities. At the same time, SOTIF standard prescribes neither the acceptable level of risk nor the effort required to reach it, neither implicitly nor explicitly. Each use case shall be considered separately. Although the required technical measures can be shared between use cases (e.g. using safety goals as proxy similar to ISO 26262), safety of each use case shall be eventually proven separately, unless the use cases considered belonging to one equivalence class (see §5.3 below).

5.1.1. Use-cases for hazard identification

A use-case consists of the functional range, desired behavior, functional system boundaries, and the scenario (cf. [4]). Scene is the static part of scenario. The class diagram of the use case according to [3] is presented on the Figure 5.

![Figure 4. Acceptable risk and required effort: ISO 26262 and SOTIF.](image)

According to ISO 26262, only the most hazardous use case is relevant for the definition of the maximum ASIL for the relevant safety goal. Then, the defined safety goal is used as a proxy for all technical activities. At the same time, SOTIF standard prescribes neither the acceptable level of risk nor the effort required to reach it, neither implicitly nor explicitly. Each use case shall be considered separately. Although the required technical measures can be shared between use cases (e.g. using safety goals as proxy similar to ISO 26262), safety of each use case shall be eventually proven separately, unless the use cases considered belonging to one equivalence class (see §5.3 below).

![Figure 5. Use-case definition according to [3].](image)
Example of a fully-defined use-case is given below:

1. **Functional range**: Automated Emergency Braking (AEB)
2. **Desired behavior**: The system activates braking if a collision with an obstacle is considered highly probable
3. **System boundaries**: Camera / radar system, AEB ECU, braking ECU, braking system.
4. **Scenario**
   4.1. **Actions and events**:
      i) The ego vehicle (i.e. the vehicle equipped with the AEB functionality considered in the analysis) is moving forward
      ii) No obstacle exists in the path of the ego vehicle within the hazardous distance (2 sec)
   4.2. **Goals & values**
      i) Goal: keep travelling down the road, no lane change, no direction change
      ii) Value: preferred speed, lateral distance, longitudinal distance
   4.3. **Scene**
      4.3.1. **Dynamic elements**: other vehicles on the road (e.g. a bus in front of the ego vehicle, at distance of 3 sec and another car behind the ego vehicle, distance 5 sec)
      4.3.2. **Scenery**
         i) Intra-urban
         ii) Right-hand traffic
         iii) Two-lane road: right lane and left lane
         iv) Bicycle lane on the right of the right lane
         v) Pedestrian pass on the right of the bicycle lane
         vi) Road is symmetric (i.e. it has two lanes, bicycle lane and pedestrian pass in the opposite direction)
     4.3.3. **Self-representation**: ego vehicle is a passenger car.

New use-cases are obtained via iterating the changeable parameters. Changeable parameters include actions and events, values, and parameters of the scene.

In order to prevent the explosion of the number of considered use cases, creativity is required in the iteration process. Only parameters relevant to the chosen functionality shall be iterated. The list above considers the AEB functionality, which influences the longitudinal dynamics of the vehicle. Therefore, lateral conditions (e.g. cars travelling in other lanes) are irrelevant for the hazards related to the AEB. They should not be iterated.

5.1.2. **Hazard evaluation**

The hazards are evaluated on the basis of their probability. SOTIF considers non-probabilistic models for system limitations, i.e. the system will experience a malfunction whenever a hazardous use case shows up. Therefore, the probability of a hazard is fully defined by the probabilistic characteristics of the hazardous use case.

Probabilities of the realization of a use case are expressed per hour of driving or per kilometer. HARA shall define this probability, which is later used for the definition of required validation mileage (see §4.3). Obviously, the probability cannot be defined on the level of detail required for the use case definition (see list in §5.1.1). Use cases may be grouped into sets where probability of exposure may be calculated for the whole set. This strategy will require the validation to be based on the whole set of use cases.

Probabilistic modelling yields relatively precise data for determination of probabilities of hazardous use cases (e.g. [5]). An example of underlying risk-generating process is shown on Figure 6. The probability of the absorbing states on the right (both “Safe driving” and “Accidents”) is known from the traffic and accident statistics. Quality of the assessment of the probabilities of driving situations depends on the precision of the modelling of the transitional process (e.g. transition rates for Markov chain modelling).
Quality-of-fit parameters (confidence intervals, likelihood etc.) may be used to understand how good the probability assessment is. The likelihood can be increased using data from previous validations. Models yield quality-of-fit estimations directly, while for other methods of the assessment those estimations shall be drawn based on assumptions.

Known use cases (i.e. use cases identified as hazardous for a given malfunction) are part of the “Data on the Intended Functionality” (see Figure 3). They are mostly not project-specific and belong to a centralized data-store.

5.2. Triggering Events Analysis
Triggering Event Analysis (TEA) is the only SOTIF-specific WP of the Generalized Safety Lifecycle. It is important to understand that a triggering event is not a use case, it is a characteristic of multiple use cases.

The TEA consists of the identification of the triggering events (TEs) and the evaluation of their acceptability (similarly to HARA).

The suggested process includes analysis of SOTIF-critical characteristics (SOTIF-CCs), similar to critical characteristics as defined in the FMEA process (e.g. [6]), related to the functionality potentially affected by hazards related to functional insufficiencies.

5.2.1. Identification of relevant TEs
The inclusion of SOTIF-CCs harmonizes the TEA process with FMEA process and allows the implementation of the TEA via any common FMEA tool. Example of SOTIF-CCs related to camera are sharpness, brightness and classification quality. The failures are identified based on the SOTIF-CCs using a HAZOP-like procedure. Those failures may be linked directly to malfunctions previously defined within the project (normally malfunctions are defined at the step of Item Definition).

The list of triggering events shall be prepared, based on the groups described in [3]: environment and location, road infrastructure, urban infrastructure, highway infrastructure, driver behavior (including reasonably foreseeable driver misuse), expected behavior of other drivers/road users, etc. The list may be pre-defined and uploaded into the FMEA tool. Then, pre-defined TEs are linked with failures, which are in turn linked to malfunctions. A failure net linking TEs to malfunctions appears.

5.2.2. Evaluation of TE acceptability
According to [3], a TE is acceptable if all use cases related to it have low exposure probability, high controllability, or risks associated to them bear low severity. Using the definitions from ISO 26262-3, in order to qualify as acceptable a TE shall be either linked only to use cases occurring far less than 1% of all driving time (e.g. an airplane landing on a highway), or the use cases easily controllable (i.e. more than 99% of drivers may resolve the hazardous situations without risks), or the use cases where there is no risk of any bodily injury.

Having the TEA and HARA prepared in a systematic way allows easy identification of acceptable and non-acceptable TEs via database tools.
Relevant triggering events are part of the “Data on the Intended Functionality” (see Figure 3). They are not project-specific and belong to a centralized data-store. The list of TEs relevant to a given project is a subset of that general list.

5.3. Validation of the Unknown Use Cases
The goal of the validation of the unknown use cases is to show that the system under development does not pose any unacceptable risk even for the use cases which were not directly considered during the development process.

The selection of the use cases to be picked for testing shall follow the distribution of the use cases in the general population. This does not necessarily include specific use cases, especially those related to the triggering events, as they should be already sorted out during verification phase.

Basic probabilities of the use cases are given in the HARA (see §5.1). The probabilities are given per hour or per kilometer. However, just inverting the number will give under 50% of confidence in safety of this use case. The full validation mileage shall be calculated considering the probabilities of the use cases to be binomially distributed [7]. Typical validation distance for a level 2 function is measured by many hundreds of thousands of kilometers, up to one million. However, the proper planning of the validation distance is more important than its length. The conditions of driving (e.g. weather, locality, intra/extra-urban driving etc.) shall be weighted according to the distribution experienced by the target vehicles. Statistics related to the road traffic gathered by local traffic authorities (e.g. so-called “Green Book” by AASHTO [8]) is the main source of information related to this.

Driving and accidents statistics are part of the “Data on the Intended Functionality” (see Figure 3). They are not project-specific and belong to a centralized data-store.

5.4. Joint Safety Assessment Report
The Safety Assessment Report (SAR) summarizes safety evidences and provides a safety argument, i.e. explanation of the findings of the safety lifecycle. The SAR is normally written in a semi-formal way. For Generalized Safety Lifecycle, the SAR is basically created according to the requirements of ISO 26262, but includes SOTIF-related evidences as well.

No specific changes in the process of creation of the SAR are required.

6. Conclusions
The main contribution of this paper is the Generalized Safety Lifecycle, integrating the lifecycles of functional safety and the safety of the intended functionality. Distribution of the requirements between WPs is shown (Table 1). Concept of a separate databank for storage of the non-project specific data on the intended functionality is introduced (§5, Figure 3). Integration of both lifecycles makes SOTIF clearer and easier to implement.

Functional safety is often perceived as “just another requirement” from the authorities or the customer. However, if implemented based on system engineering methodology, both FS and SOTIF help improving the understanding of the system under development and making it safer. They facilitate future updates and modifications, thus reducing the cost of development and aftermarket activities.

Acknowledgments
The authors want to express their gratitude to Mr. Joseph Miller (J D Miller Consulting, formerly ZF Friedrichshafen), Mr. Konstantin Byakov, PhD (Bauman MSTU), Dr. Brian Murray, Dr. Stefan Rinkenauer and Dr. Sari Buelent (all ZF Friedrichshafen), and many other colleagues who contributed to the development of this concept.

References
[1] EAEC Technical Regulation TR TS 018/2011 Safety of Wheeled Vehicles (Moscow)
[2] ISO 26262-1:2011. Road vehicles – Functional safety (Geneva)
[3] ISO/PRF PAS 21448 2018 Road vehicles – Safety of the intended functionality (Geneva)
[4] Burton S Gauerhof L and Heinzemmann C 2017 *Making the Case for Safety of Machine Learning in Highly Automated Driving* SAFECOMP Workshops LNCS 10489 pp 5–16

[5] Lurie O and Miller J 2018 Hazard Analysis and Risk Assessment beyond ISO 26262 Management of Complexity Via Parametrization *SAE Technical Paper 2018-01-1067* information on https://doi.org/10.4271/2018-01-1067

[6] AIAG VDA 2017 *Failure Modes and Effect Analysis Manual* (Berlin)

[7] Koester F Form T Lemmer T Plaettner J 2016 Wie gut müssen automatisierte Fahrzeuge fahren – PEGASUS Pegasus Project

[8] American Association of State Highway and Transportation Officials 2001 *A Policy on Geometric Design of Highways and Streets* (Washington)