Safety of the Intended Functionality of External Human Interfaces: Gaps and Research Agenda

Manabu Okada
TIER IV, Inc.
Jacom Building, 1-12-10 Kitashinagawa,
Shinagawa-ku,Tokyo, Japan
Email: manabu.okada@tier4.jp

Barbara Gallina
Mälardalen University,
P.O. Box 883, SE-72123
Västerås, Sweden
Email: barbara.gallina@mdu.se

***Abstract*---Next-generation human-computer interactions comprise the interactions that are expected to take place between e.g., pedestrians and the so-called External Human-Machine Interfaces (EHMIs). EHMIs are expected to replace the in vehicle Driver-oriented Driver-Machine Interface in the context of high automation levels, where the driving task is fully automated. In this context, it is paramount to ensure absence of unreasonable risk due to hazards resulting from functional insufficiencies of the intended functionality or its implementation, i.e., Safety Of The Intended Functionality (SOTIF), defined in ISO 21448:2022. ISO 21448:2022 has been applied to ensure the Safety Of The Intended Driver-Machine Interface-Functionality. SOTIF does not include any specific consideration to address the so-called External Human-Machine Interface. Hence, in this position paper, we posit that the functional insufficiencies shall also embrace the socio-technical issues that might emerge in the extended automated driving system's ecosystem. These issues include faulty interactions with pedestrians, traffic-lights, and more broadly, by taking a vehicle to everything perspective (VtoX), with everything.**

Keywords---Driving Automation Levels, External Human Machine Interface (EHMI), Functional Safety (FuSa), Safety of the Intended Functionality (SOTIF), Socio-technical issues, Vehicle to everything (VtoX), Vehicle to Infrastructure (VtoI)

I. Introduction

Next-generation human-computer interactions comprise the interactions that are expected to take place between not only the driver and the in-vehicle Driver-Machine Interface (DMI) but also the interactions between e.g., pedestrians and the so-called External Human-Machine Interface (EHMI). In the context of high automation levels, where the driving task is fully automated, this EHMI is expected to replace the in vehicle Driver-oriented DMI. There are high expectations and societal and economic value of higher automated driving, but it is increasingly acknowledged that interactions with external human road users constitute a key challenge [1]. In this context, it is paramount to ensure absence of unreasonable risk due to hazards resulting from functional insufficiency of the intended functionality or its implementation, i.e., Safety Of The Intended Functionality (SOTIF), defined in ISO 21448:2022 [2]. ISO 21448:2022 fills in the gaps left by ISO 26262:2018 [3], the automotive functional safety standard that targets the absence of unreasonable risk due to hazards caused by malfunctioning behaviour of E/E (electrical and/or electronic) systems. SOTIF is applicable to intended functionalities of automated driving systems having driving automation level [4] from 1 to 5. Recently observed accidents, caused by automated driving vehicles, were due to functional insufficiency. Hence, applying SOTIF recommendations would effectively lead to achieve absence of unreasonable risk and ultimately contribute to accidents prevention. The SOTIF's cause and effect model defines that triggering conditions activate functional insufficiency. The triggering condition is a specific condition of a scenario that contributes to a hazardous behavior or an inability of human risk avoidance capability including the Safety Of The Intended Driver-Machine Interface-Functionality. For instance, in [5], authors design and execute a case study with the purpose of verifying the modeling and analysis capabilities of their framework for capturing system risks caused by new augmented reality-related dependability threats in AR-enriched DMIs. In the context of a panel organised at the annual meeting of the Human Factors and Ergonomics Society (2019), the challenges and the opportunities in developing next generation in vehicle DMI systems were discussed. In [6], authors report about a large-scale participatory design, focusing on the DMI interface of the next generation of autopilot. In [7], the authors explores road traffic interactions between highly automated vehicles and vulnerable road users (VRU). However, they does not develop the exploration in the context of the SOTIF standard. To the best of our knowledge, SOTIF does not include any specific consideration to address the so-called EHMI and no specific considerations have been formulated to address the EHMI from a SOTIF perspective. Hence, in this position paper, we posit that the functional insufficiencies shall also embrace the socio-technical issues that might emerge in the extended automated driving system's ecosystem, including faulty interactions (foreseeable mis-interactions) with pedestrians, traffic-lights, and more broadly, by taking a vehicle to everything perspective (VtoX). We highlight two gaps that we identified within ISO 21448:2022 and propose corresponding research directions or potential ways forward for ensuring the safety of the intended functionality for the V2X.

The rest of the paper is organised as follows. In Section II, we recall essential background information. In Section III, we present our proposal aimed at extending SOTIF in the context of the extended Automated Driving System's ecosystem. In Section IV, we discuss related work. Finally, in Section V, we present some concluding remarks and future work.

II. Background

In this section, we recall essential background information on: 1) next generation HMIs in the automotive context, focusing on External HMIs; 2) on-road motor vehicle Automated Driving Systems (ADSs) and 3) SOTIF.

A. External HMIs for ADS

External HMIs are interfaces conceived to enable communication between the automated driving vehicle and the Other Road Users (ORU), including Vulnerable Road Users (VRUs) [8], as well as X.

EHMIs shall be carefully conceived since studies let emerge that ORU (specifically, pedestrians) are uncomfortable when losing the possibility to communicate with a driver. Different EHMI-concepts have been proposed in literature to address human factors-related challenges. For example, some concepts are proposed in [9], [10]. These EHMI-concepts utilize the display on the vehicle facing the external environment with the purpose of representing the automated driving system's intention to the external humans. These EHMI concepts aim at addressing the socio-technical challenges inherent in the interaction between highly automated vehicles and humans in their shared environment, particularly focusing on improving safety and understanding situations where traditional forms of human-vehicle interaction might be insufficient or ambiguous. In [11], authors systematically categorize some of these EHMI-concepts by providing a taxonomy, which considers different dimensions (modality of communication i.e., visual, auditory, etc., placement of the HMI, etc.). The taxonomy proposed does not consider safety concepts such as SOTIF, but the authors acknowledge the need to further investigate aspects related to safety. In [7], the author lists 19 road traffic interaction/behavioral studies between highly automated vehicles and VRUs. The author reviews 11 empirical studies on crossing-path scenarios between highly automated vehicles and VRUs. Also, the author introduces 2 interactive behaviors in merging scenarios. Through that literature review, the author concludes that the system (for higher driving automation levels) might need to address a wide range of interactive behaviors and communication. In addition to this, the author also states that it might also be effective to add HMI to compensate for the loss of social driver-centric cues or to signal other kinds of information. Even though there is a wealth of research on external HMI that points to their importance and role they play to ensure safety in the context of ADS, there is no unified framework that bring these considerations together and provides an actionable blueprint. [1].

B. SAE Driving Automation Levels

SAE J3016 [4] published a taxonomy and definitions for terms related to on-road motor vehicle Automated Driving Systems (ADSs), which perform part or all of the Dynamic Driving Task (DDT). Within the SAE taxonomy, six levels of automation are defined. To make the paper self-contained, we recall the definitions of the terms used in this paper in what follows.

The *dynamic driving task (DDT)* is defined as all of the real-time operational and tactical functions required to operate a vehicle in on-road traffic, excluding the strategic functions such as trip scheduling and selection of destinations and waypoints. *Object and Event Detection and Response (OEDR)* refers to the appropriate handling of external situations that a vehicle encounters. This includes monitoring the driving environment (detecting, recognizing, and classifying objects and events and preparing to respond as needed) and executing an appropriate response to such objects and events (i.e., as needed to complete the DDT and/or DDT fallback). OEDR is also included as a subtask of the DDT. OEDR plays a fundamental role in executing the DDT [8]. The *Operational Design Domain(ODD)* defines specific conditions under which a particular ADS or its features are intended to operate. In [4], the routine/normal ADS operation is defined as the ADS operating the vehicle within a given ODD while there is no system failure related to DDT performance.

Table I. Brief summary of the SAE automation levels.

Level	Name	OEDR	DDT fallback	ODD
0	No driving automation	Driver	Driver	Not applicable
1	Driver assistance	Driver	Driver	Limited
2	Partial driving automation	Driver	Driver	Limited
3	Conditional driving automation	System	Fallback-ready user	Limited
4	High driving automation	System	System	Limited
5	Full driving automation	System	System	Unlimited

As shown in the Table I, at level 4 and 5, the ADS (and not the human) is fully responsible for OEDR and DDT.

- Automation level-3 is defined as "Conditional driving automation". The ADS performs the DDT, excluding DDT fallback to fallback-ready users. To promptly perform the DDT fallback, the ADS is supposed to alert and request to intervene the fallback-ready user (who regains the driver role). Mode confusion and automation surprise are analyzed, focusing on driver-oriented human-machine interface.
- Automation level-4 is defined as "High driving automation". The ADS performs the DDT, including DDT fallback. There is no human driver, nor request to intervene to fallback.
- Automation level-5 is defined as "High driving automation". The only difference with level4 is "Unlimited ODD" [4].

In both levels 4 and 5, the ADS is capable of achieving a minimal risk condition if necessary (i.e., due to out of the ODD).

C. SOTIF

As mentioned in the introduction and as defined in ISO 21448:2022 [2], SOTIF stands for Safety of the intended functionality. ISO 21448:2022 provides requirements to ensure SOTIF. These requirements augment the safety life-cycle prescribed within ISO 26262:2018 [3]. ISO 21448:2022 covers all levels of the SAE Driving Automation Levels. In what follows, we provide essential information regarding the ISO 21448:2022 clauses that we think would benefit from additional research.

ISO 21448:2022 Clause3 defines misuse as *the use of a system by humans in a manner not intended by the manufacturer or service provider.* This obviously does not include humans outside the vehicle since external humans do not use the system. In the note section of misuse, direct misuse is defined *which could be a cause for the occurrence of a hazardous behaviour of the system, is considered to be a potential triggering condition.*

ISO 21448:2022 Clause5, called *specification and design*, requires the provision of the specification and design of the system, including ODD, HMI, relevant environmental conditions, data exchange with inter-vehicle or external sources, system architecture, and mechanism supporting risk mitigation during operation. It also covers reasonably foreseeable misuse.

ISO 21448:2022 Clause7, called *Identification and evaluation of potential functional insufficiencies and potential triggering conditions*, requires the analysis to identify causes leading to hazardous behaviors. In addition to this analysis, responses of the system shall be evaluated for SOTIF acceptability. The scope of the analysis includes both system behavior, such as sensing, planning, and actions, and human factor, such as reasonably foreseeable direct or indirect misuse. The scenarios with the identified triggering conditions are evaluated to assess whether SOTIF can be achievable.

III. Gaps and Research Agenda

In this section, we highlight two gaps that we identified within ISO 21448:2022. For each gap, we propose corresponding research directions or potential ways forward for ensuring the SOTIF for the V2X ecosystem by expanding the methodology and activities in ISO 21448:2022.

A. Gap1: Problem space gaps

We found the *problem space gaps* in ISO 21448:2022 clause 5 *Specification and design*. Specifically, this clause exhibits limitations when it comes to capturing road-sharing conflicts and interactive behavior. Within this clause, there are statements that state to start SOTIF activity mainly by focusing on on-board systems, related on-board systems, and interfaces with infrastructures. However, there is no statement concerning the need to focus on the interaction between the ADS and ORUs and the fact that ADS is expected to interactively maneuver the vehicle in response to road user situations and reactions. There is no statement regarding the need to have a comprehensive road structure design, where how each road user behaves is explicitly designed. These identified limitations are due to the fact that ISO 21448:2022 is written with an ego vehicle-centered perspective and not with a space-sharing, conflicts-centered or interactive behavior-centered one.

Research Agenda1: Which additional activities would be needed to capture the intended functionality of V2X ecosystem (extension of ISO 21448:2022 clause 5)?

To address this unexplored Gap1, we discuss one hypothesis. **Hypothesis**: Selecting an appropriate problem space (when you model and analyze the space sharing conflicts and unsafe interaction behavior with the V2X ecosystem, including EHMI) could be effective in capturing the intended functionality.

For example, if the problem space were an ADS'perspective in Figure 1 Option A, it might be difficult to identify those scenarios involving many road users with interactive behavior. However, if the problem space were a road structure (e.g., intersection)'s perspective in Figure 1 Option B, we could usefully treat these scenarios as space-sharing conflicts and/or interactive behavior scenarios in the intersection. This hypothesis would also solve Gap2, which will be touched on in the following.

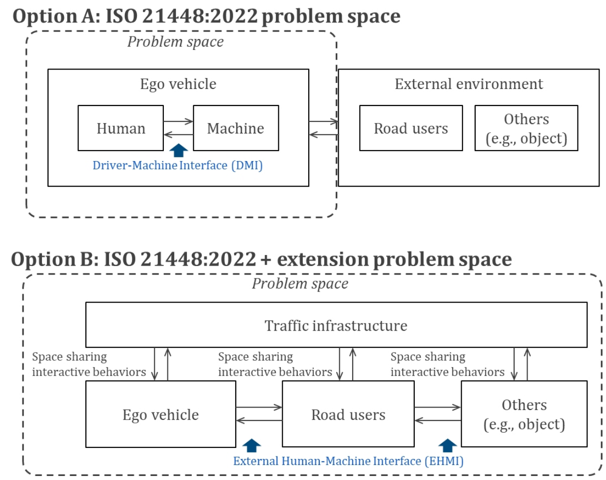

Figure 1. Selection of an appropriate problem space

Verification of Hypothesis: problem space To verify our hypothesis, we observed and analyzed space sharing conflicts related to ADS development with various problem spaces. We found that performing the additional activity of selecting an appropriate problem space enables us to better model and analyze the space sharing conflicts and unsafe interactive behavior of the V2X ecosystem with automated driving systems having driving automation level 4.

We tested our hypothesis by selecting appropriate problem spaces for an ADS and ADS-related accident and near-miss behavior as shown in Table II, based on Figure 1. We found that the causes behind them were quite diverse. We paid particular attention to selecting the problem space for each of the items in Table II. This turned out to be appropriate for identifying several causal scenarios that involve a combination of factors [12], including functional insufficiency. For each of the items in Table II, the corresponding problem space and examples of potential causal scenarios were derived as follows.

Table II. Analysis of problem space in Level 4 Driving Automation Accidents and near-miss behaviour

No	Accident [A], Near-miss [N] behaviour and Other incidents [O]	Problem space
1	[A] A woman was injured after being hit by a hit-and-run driver and was trapped under an ADS (USA, 2023) [13] [14]	Ego vehicle to X
2	[A] An ADS (golf cart) failed to detect an unattended bicycle and collided with it (Japan,2023) [15].	Ego vehicle to X
3	[A] An automated shuttle collided with a person (Japan,2021) [16].	Ego vehicle to X
4	[N] ADS got stuck at intersection (USA,2023) [17]	Road structure
5	[N] An ADS ran thorough closely with a pedestrian (USA,2023) [18]	Road structure
6	[O] An ADS drove away despite a police officer's stop(USA,2023) [19]	Ego vehicle to X

"Ego vehicle to X" is the appropriate problem space to consider to analyze No.1, No.2 and No.3 in Table II.

1) No.1: Potential causal scenarios can be insufficiency of specification that ADS could inaccurately determine a collision type, that ADS could incorrectly select the lane top maneuver instead of an emergency stop. Moreover, the ADS could fail to determine the location of the pedestrian, while the pedestrian is underneath

the vehicle. Also, ADS could fail considering an acceptable stopping location, if the lane edge is mislabeled in the semantic map. This is not a space sharing conflict nor an interactive behavior. [14]

2) No.2: One of the potential cause is insufficiency of specification that ADS could not recognize the bicycle or bicycle part and avoid a collision. This is not space sharing conflicts, nor interactive behaviours.
3) No.3: One of the potential cause is misuse. The manager could give permission to proceed, thinking the pedestrian would stop [16]. Or, inadequate specification where the ADS ignores a person's permission to drive and stops when a person is expected to be in the trajectory (Insufficiency of specification). This is not space sharing conflicts but a interactive behaviour.
4) No.6: This case may also involve several causal factors (e.g., the ADS does not feature an external human-machine interface, the ADS does not interpret the gesture of an external human, etc.). This item shows that we may embrace the socio-technical issues that might emerge depending on the ADS's capability. This is not a sharing conflict but an interactive behaviour.

"Road structure" is the appropriate problem space to consider to analyze No.4 and No.5 in Table II.

- No.4: The potential causal scenarios may involve several causal factors (e.g., possible time constraints in the traffic signal, traffic conditions and locations in the intersection, predicting the trajectory and/or behavior of other road users, including other automated driving vehicles etc,.). This includes the space sharing conflicts and interactive behaviours.

- No.5: This shows a trade-off scenario between a time-to-wait for a pedestrian to finish crossing and a time-to-exit the intersection for an ADS after the pedestrian has finished crossing within traffic signal interval. In the video [18], we can observe;
  1) An ADS stops in front of pedestrians;
  2) the ADS approaches the pedestrians even though they are in the process of crossing;
  3) The ADS begins to maneuver aggressively toward the pedestrian, taking a slight detour;
  4) The ADS moves away from the side of the pedestrian;
  5) The pedestrian notices the near miss and looks at the ADS in surprise or fear.

  This scenario is one of the difficult scenarios with EHMI within traffic time constraints. The aggressive driving captured by the video may not be acceptable in a society.

This observation particularly highlights that if we are in *Ego vehicle problem space*, the efforts to identify "SOTIF cause and effect model" are effective to lead to avoid or reduce or mitigate the risks such as accidents or near-miss behaviour due to ego vehicle triggered. However, if we are in *Road structure problem space*, we can view ego vehicle as one of the road users in the road structure such as intersection, then enable us to identify space sharing conflicts and interactive behaviour of V2X's ecosystem including VRUs. We propose to select an appropriate problem space when we model and analyze the space sharing and interaction behavior with the V2X ecosystem, including EHMI, as an extension of ISO 21448:2022.

B. *Gap2: triggering condition gaps*

We found the *triggering condition gaps* in ISO 21448:2022 clause 7 *Identification and evaluation of potential functional insufficiencies and triggering conditions*. The ISO 21448:2022 clause 7 has limitation to capture triggering conditions which could be a cause of space sharing conflicts and unsafe interactive behavior of V2X ecosystem. Most studies have focused on the behaviors of drivers and occupants in their vehicles and not on humans outside the vehicle in terms of definition of misuse in clause 3. However, for higher levels of automated driving, such as levels 4 and 5, the interactions between the vehicle and the VtoX ecosystem, including the external human, are critical to consider since external human *does not use* automated driving system, but *does use* space in a road.

This gap2 can lead to space sharing conflicts or unsafe interactive behavior. In clause 7, there are statements to identify the triggering condition of system function (e.g., sense, act, plan) and misuses of humans, who use the system. However, there are no statements to identify triggering conditions of the V2X ecosystem, especially for external humans. This could cause space-sharing conflicts and unsafe interactive behaviors. To address this Gap2 based on Gap1, we discuss one research agenda and one hypothesis.

Research Agenda2: Which methods could identify the triggering conditions related to space sharing conflicts and unsafe interactive behaviours (as an extension of ISO 21448:2022 clause 7)?

Hypothesis: To analyze the chosen problem space (refer to Figure 1), we can effectively identify causal scenarios by utilizing three view points (space, time, semantic). These scenarios include triggering conditions that potentially link to space-sharing conflicts and unsafe interactive behaviors, capturing the intended functionality.

Verification of Hypothesis: To address this Gap2, we carried out observations to better understand presence of a collision course itself to define the pivotal concept of the traffic conflict, as defined in [1]. Then we discussed the effectiveness of our hypothesis. Based on several observations, we opted for cooperative human driving to identify triggering conditions related to space sharing conflicts and unsafe interactive behaviours for cooperative driving automation(CDA) [8] as a verification of the hypothesis.

Importantly, we saw "supporting observability, predictability, and directability (OPD) between agents" as a basis for interaction and communication [7] in a youtube video [20] which are not able to cover the ISO 21448:2022 clause 7.

1) Scene1: Ego vehicle-A drives in the lane. There are many cars driving in the opposite direction, forming a cue. Objects beyond are occluded by the cue.
2) Scene2: Ego vehicle-A approaches a crosswalk but has limited view due to the cue formed on the vehicle's right-hand side.
3) Scene3: Meanwhile, vehicle-B is in the opposite lane, and the driver of vehicle-B is turning on his(her) headlights on and off.

4) Scene4: The driver of ego vehicle-A notices the actions and (s)he reduces the speed before approaching the crossroad.
5) Scene5: The next moment, a pedestrian with a stroller appears from the occlusion caused by the cue of cars.
6) Scene6: Ego vehicle-A stops beforehand and successfully avoids a collision with the pedestrian.

We based our analysis on the problem space as a road structure in Figure 1, such as a crossroad without a traffic signal. The scenario is categorized as *Crossing paths* in the five prototypical space-sharing scenarios [1].

We found differences between the means of space sharing and communication among road users could cause unsafe interacting behaviors. The behaviors can boil down to how external humans interpret communication. The existing ISO 21448:2022 clause 7 does not cover these interacting behaviors. We believed that utilizing three viewpoints (space, time, semantic) to analyze the chosen problem space (refer to Figure 1) can effectively identify causal scenarios, including triggering conditions.

We also formulated the relationships of means of space sharing and communication among road users in Figure 2. The video [20] are categorized a case2+ in Figure 2 which is a variation of case 2. Those case2 and case 3 include the EHMI and an important scope of considering the SOTIF of V2X.

In the case2+ in Figure 2 (equal to the scenario of the video), the main scope of analysis is the scenario that there are signs of intended space sharing(e.g., road and crossroad), implicit traffic control among road users(e.g., without traffic lights) and communication to control space sharing. The communications should be done among road users, including implicit and explicit communication.

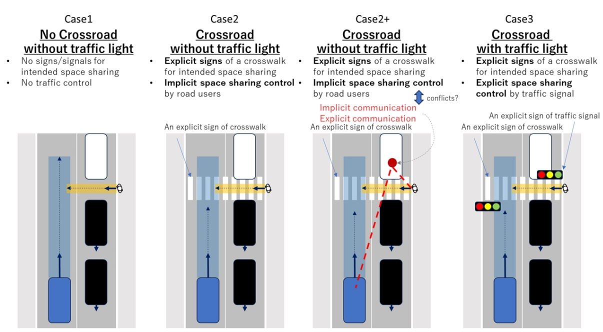

Figure 2. Explicit vs Implicit signals/signs, controls and communications

These causal scenarios, including triggering conditions, include but are not limited to the following: The actions of a pedestrian in scene 5 are not covered by ISO 21448:2022 but are covered by our hypothesis. We importantly found a *extended* triggering condition which is necessary to consider in the cooperative driving automation. The behaviours of a pedestrian with a stroller are not covered by ISO 21448:2022. The behaviours include *extended* direct misuse of human in out of the vehicle. The *extended* direct misuse is completely different from the one defined in ISO 21448:2022. However, this behavior could cause a space-sharing conflict or an accident; we have to consider or deal with it in cooperative driving automation.

Unsafe behaviour - In scene5, a pedestrian with a stroller did not confirm whether or not the driver of the ego vehicle-A had recognized him/her before he/she walking the last half of the crossroad.

Casual scenarios of unsafe behavior include triggering conditions: Semantic aspect1: Space-conflicts: A pedestrian did not think about the implicit risks, which include ego vehicle-A is coming toward him/her, because he/she has focused on other things(e.g., greeting to the vehicle-B) or needed to hurry to walk through the crossroad or have no knowledge or experiences of the similar situation. Semantic aspect2: Unsafe interactive behaviour: A pedestrian believed turning on vehicle-B's headlights on and off means "You can walk through the crosswalk"(semantic) "Now you can walk through"(time).

IV. Related Work

In the literature, various research works have addressed External Human-Machine Interfaces. In [21], authors introduce the stages of the CAVs (connected autonomous vehicles)-VRU interaction process. However, they do not focus on SOTIF, nor on triggering conditions, which could be a cause of space sharing conflicts and unsafe interactive behavior of V2X ecosystem. [22] introduces multiple studies between safety and cybersecurity in the context of V2X. We agree with this holistic approach in terms of implementation of both ISO 26262:2018 and ISO/SAE 21434:2021. Regarding holistic approaches for V2X, the authors also mention at least two different safety engineering approaches that could be considered, Holistic Single System Safety Engineering Approach and Modular Engineering Approach [23]. Our two hypotheses can be used in these studies mostly based on the activities at implementation phase of the intended functionality.

Regarding SOTIF, in [24], the authors investigate and summarize the core studies on the SOTIF of automated driving vehicles. However, they do not dig into considerations regarding forseeable misinteractions with ORU or X. In [1], the authors define space-sharing conflicts and interaction behaviour, which can play an important role to consider in the external human-machine interface. In addition to that, they also introduce their studies of implicit and explicit communication. However they do not mention how they model it and analyze it. In [25], the author states that while conducting hazards analysis and risk assessment in the context of ISO 21448:2022, mental harm shall be taken into consideration.

V. Conclusion and Future Work

The testing and deployment of Automated Driving Systems have revealed a series of insufficiencies in the intended functionality, with a special focus on external human-machine interfaces. Since the current published version of the SOTIF standard does not deal with those insufficiencies in terms of external human interaction, such as space-sharing conflicts and unsafe interactive behavior, to extend SOTIF, in this position paper, we have presented gaps and our research agenda.

Acknowledgments: This work has benefited from the results obtained within two projects: 1) a project subsidized by the New Energy and Industrial Technology Development Organization (NEDO), Japan; 2) ImmerSafe project, funded by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 764951.

References

[1] Gustav Markkula et al., ``Defining interactions: A conceptual framework for understanding interactive behaviour in human and automated road traffic," *Theor. Issues Ergon.*, vol. 21, no. 6, pp. 728–752, Nov. 2020.

[2] ISO/TC 22/SC 32, *ISO 21448:2022 Road Vehicles -- Safety of the intended functionality*, International Organization for Standardization, 2022.

[3] ISO/TC 22/SC 32, *ISO 26262:2018 Road vehicles -- Functional safety*, International Organization for Standardization, 2018.

[4] SAE J3016, *Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems*, 2021.

[5] S. Sheikh Bahaei, B. Gallina, and M. Vidović, ``A case study for risk assessment in AR-equipped socio-technical systems," *Journal of Systems Architecture*, vol. 119, no. 102250, Oct. 2021.

[6] Zhang, Ning and Jiang, Ao, ``Co-designing the next generation automatic driving vehicle HMI interface with lead-users," in *HCI in Mobility, Transport, and Automotive Systems*, ser. Lecture notes in computer science, Cham: Springer International Publishing, 2022, pp. 231–243.

[7] V. Fabricius, A. Habibovic, D. Rizgary, J. Andersson, and P. Wärnestål, ``Interactions between heavy trucks and vulnerable road users—a systematic review to inform the interactive capabilities of highly automated trucks," *Frontiers in Robotics and AI*, vol. 9, Mar. 2022, issn: 2296-9144.

[8] Automated Vehicle Safety Consortium, *AVSC Best Practice for Interactions Between ADS-DVs and Vulnerable Road Users (VRUs)*, 2022.

[9] Nissan, *Nissan IDS CONCEPT*, [28th.Oct,2018], [Accessed: January 30, 2024]. [Online]. Available: https://global.nissannews.com/ja-JP/releases/print/151028-01-e?source=nng.

[10] Volkswagen, *Light staging and exterior HMI - tiguan -- visual modality*, 17th.Oct,2018, [Accessed: January 30, 2024]. [Online]. Available: https://www.volkswagen-newsroom.com/en/evolution-of-light-4261/light-staging-and-exterior-hmi-tiguan-visual-modality-4266.

[11] Debargha Dey et al., ``Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces," *Transportation Research Interdisciplinary Perspectives*, vol. 7, p. 100 174, 2020, issn: 2590-1982. doi: https://doi.org/10.1016/j.trip.2020.100174.

[12] Leveson, Nancy G. and Thomas, John P., *STPA handbook*, MIT Partnership for Systems Approaches to Safety and Security (PSASS), Cambridge, Massachusetts, U.S., 2018.

[13] abc7NEWS, *Woman injured after being struck by sf hit-and-run driver, trapped under autonomous car, cruise says*, 4th.Oct, 2023, [Accessed: February 16, 2024]. [Online]. Available: https://abc7news.com/san-francisco-woman-injured-pedestrian-crash-cruise/13857047/.

[14] Quinn Emanuel Trial Lawyers, ``Report to the boards of directors of cruise LLC, Gm cruise holdings LLC, and general motors Holdings LLC regarding the October 2, 2023 accident in San Francisco," Quinn Emanuel Trial Lawyers, Report, 2024.

[15] NHK, *Level 4 automatic driving accident: cameras failed to recognize bicycles.* 10th.Nov,2023, [Accessed: February 16, 2024]. [Online]. Available: https://www3.nhk.or.jp/news/html/20231110/k10014254121000.html.

[16] Asahishinbun-DIGITAL, *Toyota employee of automatic operator sentenced for fender bender at para athletes' village*, 6th Jan, 2022, [Accessed: February 16, 2024]. [Online]. Available: https://www.asahi.com/articles/ASQ163QJ0Q16UTIL00G.html.

[17] Jay Latta, *Brilliant demo why the current dev around self-driving vehicles is useless BS, here by Cruise.* Aug, 2023 [Online].Available: Aug, 2023, [Accessed: February 16, 2024]. [Online]. Available: https://www.linkedin.com/posts/jaylatta_brilliant-demo-why-the-current-dev-around-activity-7099036934458470400-2pFs/.

[18] r/sanfrancisco, *Cruise in front of me yesterday illegally went through a Stop sign and nearly ran over two moms and their kids.* Aug, 2023, [Accessed: February 16, 2024]. [Online]. Available: https://www.reddit.com/r/sanfrancisco/comments/15qzdrs/cruise_in_front_of_me_yesterday_illegally_went/.

[19] Levin, Tim, ``Cops pull over driverless taxi in san francisco, hilarity ensues," *Business Insider*, Apr. 2022. [Online]. Available: https://www.reddit.com/r/sanfrancisco/comments/15qzdrs/cruise_in_front_of_me_yesterday_illegally_went/?rdt=59983.

[20] KeyakiMatuya, *Purius-nopassing-masakano-hitodasuke*, Youtube, 11st.Jul,2018, [Accessed: 2024-1-30]. [Online]. Available: https://www.youtube.com/watch?v=R-8qbeuEUwQ.

[21] Reyes-Muñoz, Angélica and Guerrero-Ibáñez, Juan, ``Vulnerable road users and connected autonomous vehicles interaction: A Survey," *Sensors*, vol. 22, no. 12, 2022. doi: 10.3390/s22124614.

[22] Autotalks, ``Infrastructure for Safety," Autotalks, Tech. Rep., 2021. [Online]. Available: https://auto-talks.com/wp-content/uploads/2021/07/V2X-Infrastructure-for-Safety-RSU.pdf.

[23] 5GAA Automotive Association, ``Safety treatment in connected and automoated driving functoins report," 5GAA Working Group 2, Tech. Rep., 2021.

[24] Hong Wang et al., ``A survey on an emerging safety challenge for autonomous vehicles: Safety of the Intended Functionality," *Engineering*, 2024, issn: 2095-8099. doi: https://doi.org/10.1016/j.eng.2023.10.011.

[25] B. Gallina, ``Safety of the Intended Functionality: What about Mental Harm?" In *CARS 2024*, Leuven, Belgium, Apr. 2024. [Online]. Available: https://hal.science/hal-04558509.