Title

Developing safety metrics for human-system interaction in heavy-duty automated vehicles

Author

Cosmin-Spanoche, Anna Sofia

Publication Date

2024

UNIVERSITY OF CALIFORNIA

Los Angeles

Developing safety metrics for human-system interaction in heavy-duty automated vehicles

A thesis submitted in partial satisfaction

of the requirements for the degree

Master of Science in Aerospace Engineering

by

Anna Sofia Cosmin-Spanoche

2024

ABSTRACT OF THE THESIS

Developing safety metrics for human-system interaction in heavy-duty automated vehicles

by

Anna Sofia Cosmin-Spanoche

Master of Science in Aerospace Engineering

University of California, Los Angeles, 2024

Professor Ali Mosleh, Chair

Incorporating automated driving technologies in commercial heavy-duty operations aims to increase traffic efficiency and safety. However, developers, fleet operators, and regulators must address the unique safety risks that Automated Driving System (ADS) technology introduce. Potential Heavy-Duty Automated Vehicle (HD-AV) fleet operations envision a team of human and machine agents, including the ADS, an onboard safety driver, a fleet operations center, and, in some cases, an onboard safety operator. The complex interactions between these human and machine agents must be addressed when determining the system's safety requirements and design. Safety metrics usually focus on ADS performance, but to adequately inform system design and safety requirements, these metrics must also focus on human-system interactions. This work proposes a system model-based approach to develop human-system interaction metrics focusing on hardware, software, and human impacts on the operational safety of automated driving system fleets.

The thesis of Anna Cosmin-Spanoche is approved.

Jiaqi Ma

Tsu-Chin Tsao

Ali Mosleh, Committee Chair

University of California, Los Angeles

2024

Table of Contents

List of Figures

List of Tables

List of Acronyms

AD	Aggressive Driving
ADAS	Advanced Driver Assistance Systems
ADS	Automated Driving System
CDL	Commercial Driver’s License
CoTA	Concurrent Task Analysis
DDT	Dynamic Driving Task
DMS	Driver Monitoring System
ESD	Event Sequence Diagram
FMCSA	Federal Motor Carrier Safety Administration
FMS	Failure Mitigation Strategy
FOC	Fleet Operations Center
FT	Fault Tree
HD-AV	Heavy-Duty Automated Vehicle
HMI	Human-Machine Interface
HRA	Human Reliability Assessment
MOC	Maintenance Operations Center
MRS	Minimal Risk Condition
NHTSA	National Highway Traffic Safety Administration
ODD	Operational Design Domain
OEDR	Object and Event Detection and Response
PRA	Probabilistic Risk Assessment
SAE	Society of Automotive Engineers
STPA	Systems Theoretic Process Analysis
TTC	Time-To-Collision
VSSA	Voluntary Safety Self-Assessment

Acknowledgement

The following conference papers have been adapted for inclusion in this thesis:

- Cosmin-Spanoche, A. C. Correa-Jullian, X. Xia, A. Mosleh and J. Ma (2024), ‘Exploring safety-related metrics to assess human-system interactions in heavy-duty automated vehicles’, Accepted in 34th European Safety and Reliability Conference (ESREL 2024), Cracow, Poland, 2024.
    - **Anna Cosmin-Spanoche:** Methodology, Visualization, Investigation, Writing - Original Draft, **Camila Correa-Jullian:** Methodology, Writing - Review & Editing, **Xin Xia:** Supervision, **Ali Mosleh, Jiaqi Ma:** Conceptualization, Funding acquisition
- Cosmin-Spanoche, A. C. Correa-Jullian, X. Xia, A. Mosleh and J. Ma (2024), ‘Deriving human-system safety metrics for heavy-duty automated vehicle applications through Concurrent Task Analysis’, Accepted in 17th International Conference on Probabilistic Safety Assessment and Management & Asian Symposium on Risk Assessment and Management (PSAM17&ASRAM2024), Sendai, Japan, 2024.
    - **Anna Cosmin-Spanoche:** Methodology, Visualization, Investigation, Writing - Original Draft, **Camila Correa-Jullian:** Methodology, Writing - Review & Editing, **Xin Xia:** Supervision, **Ali Mosleh, Jiaqi Ma:** Conceptualization, Funding acquisition

This research is partly supported by the University of California Institute of Transportation Studies.

Introduction

Automated driving technology has received significant attention in a variety of transportation-related applications, including passenger transport, onboard driver assistance features, and commercial applications Automated driving systems (ADS) and their associated technologies possess significant implications for transportation, commercial sectors, and legal frameworks, and the degree to which they are adopted in the future resides heavily on in-depth identification and response to the risks associated with them. The Society of Automotive Engineers (SAE) categorizes automation into six levels, ranging from 0 to 5, based on the degree of human versus automated agent involvement in the Dynamic Driving Task (DDT). At Levels 0 to 2, the human driver is primarily responsible for performing the DDT, albeit with support from driving assistance features. For Levels 3 and 4, the ADS assumes responsibility for the DDT within a defined Operational Design Domain (ODD), with varying degrees of fallback requirements from a human operator. At Level 5, the ADS is capable of performing all DDTs and fallback functions independently, without restrictions to a specific ODD (SAE International, 2021).

The integration of Automated Driving Systems (ADS) technology into heavy-duty commercial transport operations is driven by the objectives of reducing traffic collisions and incidents attributable to human errors, while enhancing efficiency and extending operational hours (Bhoopalam et al., 2023). Heavy-duty vehicles, defined as those weighing over 26,001 lbs., encompass buses, construction vehicles, trucks, and other applications. Predominantly, heavy-duty vehicles are utilized in commercial trucking operations, which account for approximately 70% of the annual freight tonnage in the United States(Talebian & Mishra, 2022). Various reports have evaluated the impact of replacing human drivers with automated technologies, estimating crash reduction rates from 50% (Shetty et al., 2022) to 90% (Bonnefon, 2021). Furthermore, the operational hours of commercial trucking are currently restricted by the Hours-of-Service regulation of the U.S. Department of Transportation (USDOT). Increased automation could potentially extend these hours towards continuous 24/7 operation. Additionally, incorporating platooning and live traffic data could enhance fuel efficiency, thereby reducing operational costs (Talebian & Mishra, 2022). Heavy-duty automated vehicle (HD-AV) operations can encompass a range of trucking applications, including middle-mile, drayage, and long-haul, which each present unique challenges, tasks, operational profiles, and safety requirements. While there are over twenty companies in various stages of HD-AV development internationally, this technology has not reached deployment on public roads yet. Therefore, efforts must be directed towards understanding and assessing the new risks that arise by increasing the number of heavy-duty automated vehicles on the road, expanding their operational hours, and the use of these emerging technologies.

HD-AV operations currently envision operations within the range of Levels 2-4 of driving automation, with many planned operations including human agents in some capacity (Bhoopalam et al., 2023). These agents may serve as fleet operators, safety drivers, or safety operators. Their roles include performing sections of driving prior to entering the ODD, conducting control transitions, and responding to emergency scenarios. Therefore, it is critical to assess the safety implications of these human-ADS interactions. To quantify the safety of HD-AV systems, a set of metrics including the unique aspects of human-system interaction must be defined. Guidelines for creating and implementing ADS safety and performance metrics are detailed in standards such as UL4600, ISO26262, and SOTIF (International Standardization Organization., 2018; International Standardization Organization, 2022; UL Standards and

Engagement, 2023).However, these do not fully consider the role of safety drivers in the current planned use cases of HD-AVs. While current discussions about safety drivers in HD-AVs are limited to testing phases before public deployment, safety drivers are likely to remain present during nominal operations due to legal and regulatory requirements. Metrics for assessing ADS performance are mostly limited to assessing the functional safety of the system itself rather than operational processes and the hybridity of human-system interactions. Traditional safety metrics, referred to as *lagging* metrics, record the occurrence of failure events such as accidents, injuries, and fatalities. Examples of these metrics include the incident rate per miles driven or number of fatalities. However, there is growing interest into developing *leading* metrics, which collect data related to safety-relevant events that do not lead to a catastrophic event. These surrogate safety metrics (SSMs), such as Time-To-Collision (TTC), Deceleration Rate to Avoid the Crash (DRAC), and DeltaV, aim to quantify interactions with other road vehicles during "near misses" and non-failure events (Wang et al., 2021). Leading metrics may provide further indications of system safety, enabling standardization and comparison of non-failure events before vehicles enter public deployment and high-severity incidents occur (Reiman & Pietikäinen, 2012).

A systematic model-based approach is needed to construct metrics related to human-system interactions in HD-AV operations. Qualitative model-based risk assessment methods such as CoTA and System-Theoretic Process Analysis (STPA) can serve as a basis to conduct hazard identification for quantitative risk models used in probabilistic risk assessment (PRA) such as Event Sequence Diagrams (ESDs) and Fault Trees (FTs). The human-system interaction in autonomy (H-SIA) method using CoTA models was initially applied to analyze autonomous maritime vessels with remote operators (Ramos et al., 2020a).

This work aims to explore potential human-system interaction and collaboration metrics in the context of HD-AV operations. It is structured as follows: a general reference fleet that covers most commercial HD-AV operations is developed, forming a system breakdown and definition of operational phases based on system interactions. The system is further decomposed into scenario-based modeling through ESDs and critical task identification through CoTA. To analyze the safety of human-system interactions, the models focus on the most critical interactions between humans and ADS, using this information to construct relevant metrics.

Heavy-Duty Vehicles and ADS

Automated driving technology is becoming increasingly common in various transportation applications worldwide, including passenger "robotaxi" transport, driver support features integrated in passenger vehicles, and commercial transport applications. The wide range of potential large-scale applications have underscored the importance of adequately identifying, characterizing, and estimating the risks introduced by this new technology in the transportation environment. The degree of driving automation across applications is varied, depending on the Dynamic Driving Task (DDT) division and allocation between human and machine agents. The DDT is defined as the operational and tactical functions required to operate a vehicle in on-road traffic, however it excludes functions such as trip scheduling and selection of destinations and waypoints. A description of DDT functions, and whether they are classified as operational or tactical are given in Table 1. Subtasks 3 and 4 are also referred to as the object and event detection and response (OEDR).

Table 1: DDT functions.

DDT Function Num.	**Description**	**Operational/Tactical**
1	Lateral vehicle motion control via steering	Operational
2	Longitudinal vehicle motion control via acceleration and deceleration	Operational
3	Monitoring the driving environment via object and event detection, recognition, classification, and response preparation	Operational and Tactical
4	Object and event response execution	Operational and Tactical
5	Maneuver planning	Tactical
6	Enhancing conspicuity via lighting, sounding the horn, signaling, gesturing, etc.	Tactical

SAE classifies the extent of a vehicle's automation into six levels, ranging from 0-5. Levels 0-2 refer to driving assistance features, where the human driver remains in charge of the DDT. Levels 3-5 gradually transfer the DDTs to the ADS (SAE International, 2021). The ADS agent is referred to as "the hardware and software that are collectively capable of performing the entire DDT on a sustained basis, regardless of whether it is limited to a specific Operational Design Domain (ODD)". Distinctions between Level 3 and Level 4 reside in that in the former, the human driver is still responsible for acting as a fallback-ready user. Similarly, Level 4 and Level 5 differ in that the operation – including DDT fallback – is not restricted to specific ODD requirements. A summary of the SAE Levels of Autonomy and their relation to DDT and OEDR tasks are given in Table 2.

Table 2: SAE autonomy levels.

Level	**Name**	**Definition**	**Lat./ Long. Vehicle Motion Control**	**OEDR**	**DDT Fallback**	**ODD**
0	No Driving Automation	The performance by the driver of the entire DDT, even when	Driver	Driver	Driver	N/A

Level	**Name**	**Definition**	**Lat./ Long. Vehicle Motion Control**	**OEDR**	**DDT Fallback**	**ODD**
		enhanced by active safety systems.				
1	Driver Assistance	The sustained and ODD-specific execution by a driving automation system of either the lateral or the longitudinal vehicle motion control subtask of the DDT (but not both simultaneously) with the expectation that the driver performs the remainder of the DDT.	Driver and System	Driver	Driver	Limited
2	Partial Driving Automation	The sustained and ODD-specific execution by a driving automation system of both the lateral and longitudinal vehicle motion control subtasks of the DDT with the expectation that the driver completes the OEDR subtask and supervises the driving automation system.	System	Driver	Driver	Limited
3	Conditional Driving Automation	The sustained and ODD-specific performance by an ADS of the entire DDT with the expectation that the DDT fallback-ready user is receptive to ADS-issued requests to intervene, as well as to DDT performance-relevant system failures in other vehicle systems and will respond appropriately.	System	System	Fallback-ready user (Becomes the driver during fallback)	Limited
4	High Driving Automation	The sustained and ODD-specific performance by an ADS of the entire DDT and DDT fallback without any expectation that a user will need to intervene.	System	System	System	Limited
5	Full Driving Automation	The sustained and unconditional (i.e., not ODD-specific) performance by an ADS of the entire DDT and DDT fallback without any expectation that a user will need to intervene.	System	System	System	Unlimited

The development of HD-AVs is a rapidly growing field, with over 20 unique companies in various stages of innovation internationally. Heavy-duty vehicles are defined as Class 7-8 vehicles, or vehicles weighing over 26,001 pounds (11,794 kg). The classes are denoted by the limitations of Class 7 being between 26,001 and 33,000 pounds (11,794 to 14,969 kg), and Class 8 being over 33,000 pounds (14,969 kg). The functions that heavy-duty vehicles perform are varied, ranging from construction vehicles (such as dump trucks and cement mixers) to garbage trucks, to large transit buses. However, their most common use is in commercial transport operations, which can be either short-haul or long-haul. These are further categorized into drayage, freight, less-than-truckload, and intermodal trucking, among other distinctions.

Motivations for the development of ADS include the potential reduction of collisions caused by human error, increases in commercial transport efficiency, and an eventual progression toward 24/7 operations. Various media sources have inferred that replacing humans with automated driving technologies would reduce over 90% of crashes (Bonnefon, 2021). However, this report notes that this estimate does not imply that the drivers are at fault in all the associated collision scenarios (Zhai et al., 2023). A more realistic estimate of crash reduction when replacing humans with automated driving technologies may, in fact, be closer to 50% (Shetty et al., 2022). Nevertheless, soon, humans will continue to interact with ADS systems as drivers, passengers, or supervisors, inside and outside of the vehicle.

In the United States, the safe operation of heavy-duty vehicles is principally overseen by NHTSA, who publish yearly safety reports detailing crash and fatality statistics. In 2021, there were almost 14 million medium- and heavy-duty vehicles registered. In this year, large trucks made 9% of all the vehicles involved in fatal traffic crashes despite only making up 5% of all registered vehicles that year. In addition, 72% of fatal crashes involving a large truck resulted in an injury or fatality to occupants of another vehicle, compared to 11% being occupants of the truck. (National Highway Traffic Safety Administration, 2023). To increase reliability and reduce incidents while also increasing operational coverage, many heavy-duty operators are exploring ADS as an option. However, efforts must be directed towards understanding and assessing the new risks that arise by introducing this new technology.

Currently, there are multiple prospective models of HD-AV operations, with levels of driving automation ranging from 2-4 depending on the specific application. Many of these contemplate interactions with human operators and safety drivers, either in monitoring roles or by completing sections of the drive. For instance, shorter-term applications such as “middle mile” operations consider safety drivers to complete sections of the driving operation prior to entering and after exiting the ODD. In addition to the safety driver, these commercial operation contexts may also consider the active participation of a fleet operations center (FOC), and, in some cases, an additional onboard safety operator. The human operators are not only expected to interact with the HD-AV during planned sections of the operation, but also to collaborate with the vehicle in emergency situations, manage planned and unplanned control transitions, and conduct post-incident procedures.

Control transitions between the human and the ADS occur when there is a limitation of the ADS capabilities at certain levels of autonomy. A control transition can allow the human to take control of the vehicle and operate it outside of its ODD, considering, for example, geographic or weather restrictions, vehicle or sensor failures, and encounters of traffic scenarios outside of the scope of the ODD. Additionally, while the vehicle is within the ODD, a control transition can allow the human to transfer control of the driving task to the ADS.

As a convention for this work, when the control is transferred from the driver to the ADS it is referred to as a "handover", while when the control is transferred from the ADS to the driver, it is referred to as a "takeover". The driver initiates a control transition through physical mechanisms, either by directly engaging with the steering wheel, brakes, or throttle pedals when or by requesting a handover through buttons installed in the steering wheel (Camila Correa-Jullian et al., 2024). For a handover or takeover event to occur, the HD-AV system design must clearly include the order and coordination of tasks needed to transfer control, and this can be modeled through a task-based analysis like CoTA.

Since human operators are involved in various stages of HD-AV operations, it is of interest to analyze the safety implications of the interaction of these human agents with the vehicles. To support claims that HD-AV systems are safe, data on these systems needs to be collected, and a set of universal metrics needs to be defined to conduct fair comparisons and safety assessment.

Metrics to Assess Human-System Interaction

Safety metrics are an effective framework to analyze overall system safety in tandem with other indicators of risk, and are used in a variety of industries, such as maritime, nuclear, and aviation. Metrics can either be considered *lagging*—based on past incidents, or *leading*—measuring non-accident scenarios that point to the likelihood of an accident (Reiman & Pietikäinen, 2012). While crash data is widely used in safety analysis, crashes are relatively rare occurrences, so it is often difficult to find meaningful patterns due to the limited data points. Thus, surrogate safety metrics (SSMs) have been developed to measure traffic conflicts, which are statistically linked to crashes, but do not necessarily result in a crash (Bin-Nun et al., 2023). In addition, there have been several SSMs developed to analyze ADS performance (Automated

Vehicle Safety Consortium, 2021; Wang et al., 2021). SSMs have been proposed to measure and contrast the driving safety performance of automated vehicles, some stemming from traffic engineering metrics created for human-driven vehicles, and others proposed following results from real-life and simulator environments. Some metrics proposed through this analysis include Time-to-Collision (TTC), which is the time until a collision between two vehicles in the scenario environment if they maintain present velocities. In addition, Aggressive Driving (AD), a binary metric, assesses whether an autonomous vehicle performs “repeated maneuvers (longitudinal and lateral accelerations) above specified thresholds completed by the ego vehicle that are defined as less safe” (Wishart et al., 2020).

When considering humans operating non-autonomous vehicles, there have been several onboard driver monitoring systems as part of Advanced Driver Assistance Systems (ADAS) developed. These systems aim to assess the degree of attention the driver has when manually driving, through physical cues and eye movements which are analyzed through computer vision (Masala & Grosso, 2014).In addition there are retroactive self-reported assessment methods, namely the driving inattention scale (ARDES), which dictates the propensity of drivers to make attention-related errors. Originally developed in Argentina, ARDES has been validated to observe cross-cultural driving inattention patterns across various countries including Spain, the UK, and the USA (Castro et al., 2024).

In addition, there has been research done with regards to takeover requests (TORs) and the metrics associated with them. These human factors studies have analyzed characteristics like attention, fatigue, and reaction time, which are all relevant for HD-AV systems. Other studies have focused on emergency situations, and the associated task load and complexity of takeover events. When providing a full view of operational scenarios, modelling emergency situations is also important and highlights the criticality of safety metrics.

To develop effective metrics to assess interactions between humans and the ADS system, analysts and developers should rely on a combination of quantitative and qualitative metrics. For instance, metrics assessing TORs are based on not only the operator's reaction time, but also the quality of the takeover. They are used to identify system and operation design elements that hinder or enhance the performance of the safety driver and the ADS (DeGuzman et al., 2021; Morales-Alvarez et al., 2020; Zhang et al., 2019). Sources for human and ADS safety metrics are shown in Figure 1.

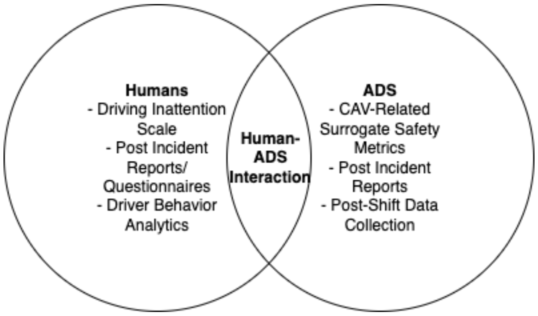

Figure 1: Depiction of metric sources for humans and ADS.

Since there has been limited research done into metrics assessing the quality of interactions between human operators and the ADS system, there is a need to explore metrics that can help to explain these interactions and compare them across systems. Driving simulators, such as OpenCDA and CARLA, have been helpful to study aspects of driving including vehicle functions and behavior, as well as human factors. They are relatively low-cost, safe, and have the potential to more adequately model potential incidents where there is limited data (Dosovitskiy et al., 2017; Xu et al., 2021, 2023). Not only do simulators allow for the collection of data regarding vehicle performance, but when experimenting with safety drivers, certain factors of the human-system interaction stage can be assessed. Several studies have observed variables like

ADS takeover decision-making, trust calibration, and organizational factors that leading to perception of ADS through data sources like physiological state and post-experiment questionnaires (Chu et al., 2023).

System Modeling and Hazard Identification

An operational safety hazard identification methodology developed for complex socio-technical systems is employed to develop new metrics for human-system interaction. This structured process combines several risk assessment techniques, including ESDs, CoTA, FTs, and STPA to analyze the interactions between agents in a complex system and identify hazards. It was initially developed to study human-system interactions in ADS operations, specifically for the case of Level 4 fleets used in Mobility as a Service (MaaS) passenger transport providers. This hazard identification method is divided into three stages: system modelling, scenario modelling, and hazard identification (Correa-Jullian et al., 2024). The present work utilizes the first stages of the hazard identification methodology for the HD-AV scenario, applying the ESD and CoTA steps to derive metrics specifically for human-ADS interactions. Stage 1 includes modelling the system by defining agents and their high-level tasks (Step 1) and defining operational phases and the transitions between them (Step 2). Stage 2 (scenario modelling) involves documenting operational phases through ESDs (Step 3) and modelling tasks and interactions using CoTA (Step 4). Stage 1 was implemented for the HD-AV system, and a list of preliminary human-interaction metrics based on the functional breakdown and operational phases were developed (Cosmin-Spanoche et al., 2024). In this work, Steps 3 and 4 from Stage 2 are implemented to identify and refine the metrics employing a structured, task-oriented approach. Further details on the method and extended results can be found in (Correa-Jullian et al., 2024). A depiction of the initial tasks utilized to derive metrics and hazard scenarios is shown in Figure 2.

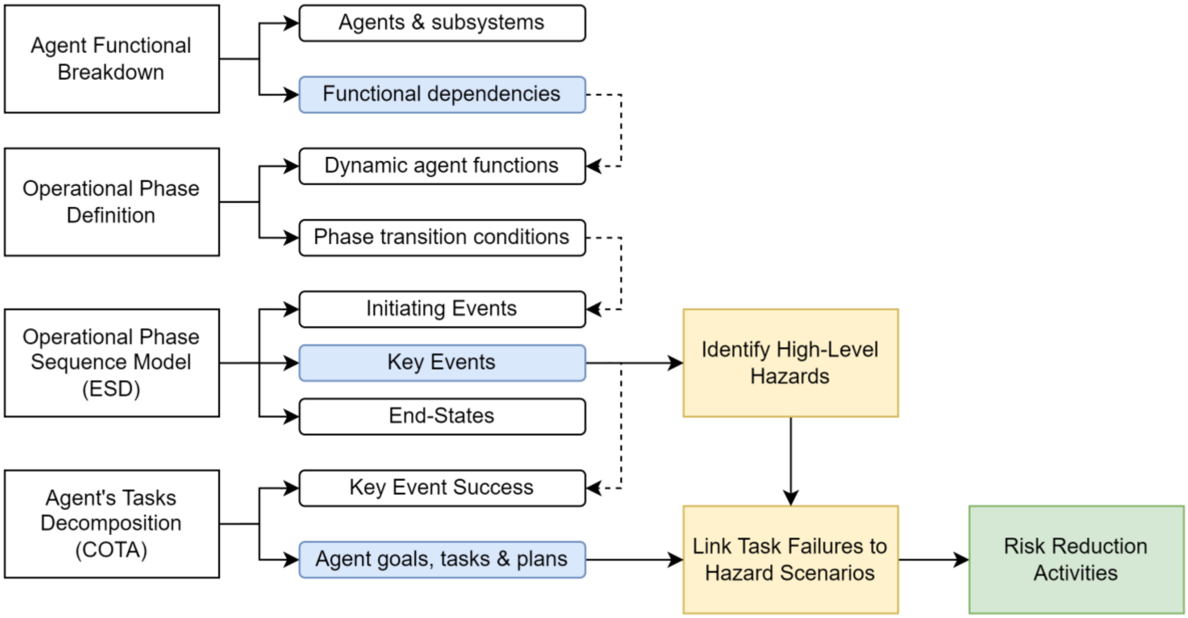

Figure 2: Operational safety hazard identification methodology.

The first two steps of the method involve modelling the system by performing a breakdown for each agent and defining operational phases. A generic model of HD-AV operations was developed by analyzing companies currently developing HD-AV technology and following NHTSA guidance for ADS system design. The relevant agents in the reference HD-AV are the ADS, safety driver, safety operator (optionally), and fleet operations center (FOC). After this, an operational system breakdown was created to model the nominal stages of operation and transitions between these stages. The tasks leading to stage transitions were further analyzed through ESDs, and the key events occurring in these ESDs were decomposed into CoTA models.

Reference Fleet

In the United States, commercial trucking corporations are primarily structured under two different paradigms. The first model is led by owner-operators, who are independent contractors that lease their services to trucking companies or directly to shippers. The second consists of fleet operators, which rely on companies that own and operate their own fleet of trucks and are responsible for hiring drivers to operate these vehicles. Operators of a commercial truck must possess a valid Commercial Driver's License (CDL) and must undergo extensive background checks and training prior to operation. The Federal Motor Carrier Safety Administration (FMCSA) regulates standards relating to commercial trucking, including creating appropriate CDL tests and enforcing regulations.

For this work, a reference fleet was defined based on a current sample of companies developing and testing in the HD-AV space. These include Aurora, Kodiak Robotics, Torc Robotics, TuSimple (which ceased operations in the United States in late 2023), and Ike (currently acquired by Nuro). Each of these companies were selected because they created and published a Voluntary Safety Self-Assessment (VSSA), and information on their safety operations was presented in their respective documents.

Characteristics such as ODD restrictions, vehicle sizing, and human-autonomy teams were selected to construct a representative model of HD-AVs in the industry. In addition, guidance from NHTSA for design of ADS systems was employed. Although many existing HD-AV companies have significant differences in their designs and processes, this general case is representative of the designs and processes of the short-term development of this industry (National Highway Traffic Safety Administration, 2017).

The reference fleet consists of retrofitted Class 8 commercial vehicles owned by fleet operators, who are responsible for developing and implementing the ADS (in coordination with an ADS developer), training safety drivers, and monitoring operations through a fleet operations center. These vehicles operate on highways and on/off ramps, which are denoted the “middle mile” for commercial goods transport, as depicted in Figure 3. A summary of characteristics of the reference fleet are given in Table 3.

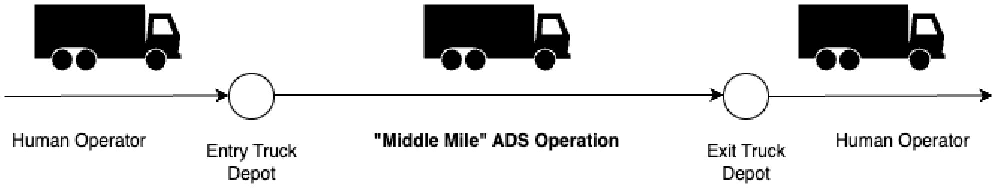

Figure 3: Depiction of ADS middle mile operations.

Table 3: Reference fleet characteristics.

Reference Fleet Characteristic	**Description**
Operational Profile	Highway and on/off ramp operations – “Middle mile” for commercial goods transport. Fleet operator owns vehicles and is responsible for training safety drivers. Pre-shift inspection checklists and special maintenance requirements for ADS components.
Vehicle Characteristics	Retrofitted Class 8 commercial vehicles (>33,000 lbs./14,969 kg)
ODD Map Restrictions	Location restricted through geofenced maps.
ODD Road & Weather Conditions	Highway roads; Fair weather; High visibility conditions; Clear to mild rain
Human-Autonomy Teams	Case 1: 1 trained onboard safety driver, 1 trained onboard safety operator, fleet operations center, ADS Case 2: 1 trained onboard safety driver, fleet operations center, ADS
ADS Capabilities	Take input from sensor data to perform DDT task, automatic detection of moving to fallback/MRC state, notify driver/operator if approaching limits of ODD.

The reference HD-AV fleet operates in a restricted ODD with location constrained by geofenced maps. The ODD is also restricted by weather conditions, with operations taking place in high visibility and fair-weather conditions, with the most severe weather condition being light rain. Although the idea of dedicated autonomous lanes has been proposed, ideally HD-AVs should be able to operate in mixed traffic scenarios, so those conditions are considered here. These vehicles operate at a nominal Level 4 of autonomy, which denotes that the vehicle itself must be able to initiate fallback autonomously if triggered by ODD exits, vehicle safety-critical failures, and other emergency situations. In the event the HD-AV exits the ODD under unplanned circumstances a DDT fallback is triggered until a Minimal Risk Condition (MRC) is achieved. In addition, operations consider the presence of a trained safety driver onboard and, in some cases, a trained safety operator, to oversee operations and intervene in the vehicle's operation for safety reasons (Aurora, 2021; Ike, 2019; Kodiak, 2020; Torc, 2021; TuSimple, 2020).

System Breakdown and High-Level Tasks

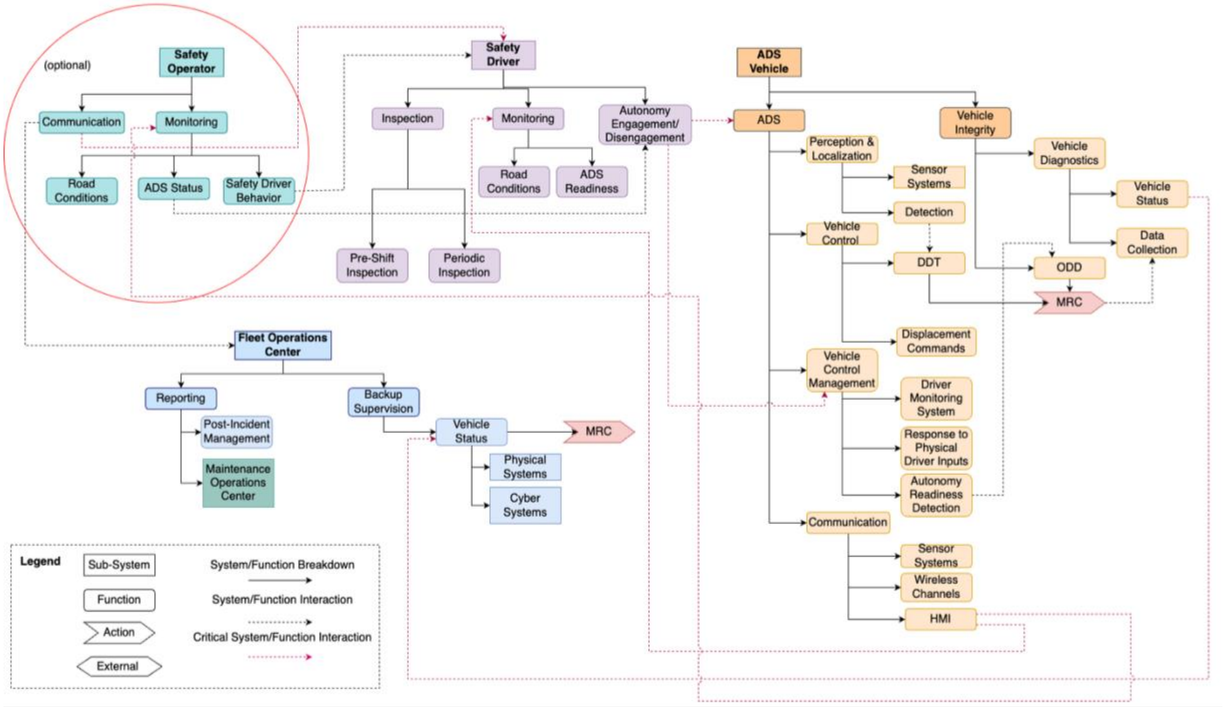

Figure 4: System functional breakdown.

To analyze the reference fleet's operations, the operations of the HD-AV system are broken down into different functional agents. The HD-AV system includes four main agents: the

ADS, safety driver, safety operator (optional), and fleet safety operator. These agents and their main task categories are defined in Figure 4 and described in depth in the following sections.

ADS

The first agent present in the HD-AV system is the ADS itself, which is the software and hardware responsible for performing the DDT within the limits of the ODD. The ADS has a nominal Level 4 of autonomy, which means that it is designed to function without the need for a human to take over the vehicle while operating within the ODD. The ADS can perform the DDT fallback to reach MRC if required. However, current operations still require a safety driver to perform sections of transit outside of the ODD (i.e. before and after middle mile operations) and provide backup in emergency cases. In addition, the ADS contains a driver monitoring system (DMS) to assess driver attention(Ayas et al., 2023). A description of the ADS high-level tasks is shown in Table 4.

Table 4: ADS task categories and descriptions.

Task Category	**Description**
Inspection	Collect and process sensor input data coming from GPS, camera, radar, and LiDAR systems.
Vehicle Control	Plan and implement the DDT (Dynamic Driving Task) while the vehicle is under computer control. Issue actuation commands to the vehicle, including steering, throttle, braking, and indicator commands. Respond to physical inputs from Safety Driver. Assume fallback state when vehicle begins to exit ODD.
Vehicle Control Management	Determine autonomy readiness based on road conditions and ODD. Monitor Safety Driver behavior through DMS. Inform Safety Driver/Operator about automated system state.
Communication	Communicate with Safety Driver and Operations Center if there is an unexpected event while ADS is activated.
Vehicle Diagnostics	Assess and report status of vehicle subsystems, both related to ADS and non-automated systems.

Safety Driver

The safety driver is a commercial vehicle operator who possesses a valid CDL and has undergone training for commercial driving operations and interactions with the built-in ADS. It is expected that the safety driver is trained in identifying the ODD requirements, ADS limitations, and is instructed on emergency procedures. Their high-level responsibilities include driving the vehicle outside its ODD and engaging and disengaging the automated driving phase. A description of the safety driver's high-level tasks is shown in Table 5.

Table 5: Safety Driver task categories and descriptions.

Task Category	**Description**
Inspection	Conduct pre-trip inspection of safety-critical vehicle systems. Inspect truck and trailer every time the truck stops.
Monitoring	Monitor road and behavior of vehicle.
Autonomy Engagement/Disengagement	Engage and disengage vehicle's autonomy system. Take control of the vehicle in case of a disengagement. Manually drive the vehicle when it is outside of its ODD.
Communication	Communicate with Operations Center about issues.

Safety Operator

The safety operator is an additional human agent onboard the vehicle in the passenger's seat whose responsibility is to monitor road conditions and the state of the HD-AV. These operators may interact with dedicated Human Machine Interface (HMI) display to identify any potential issues internally with the ADS or externally in order to warn the safety driver. In addition, they serve as a party that enables communication between the safety driver and fleet operations center. In the event the fleet operations do not include a safety operator, the safety operator's tasks are incorporated into the safety driver's responsibilities. A description of the safety operator's high-level tasks is shown in Table 6.

Table 6: Safety Operator task categories and descriptions.

Task Category	**Description**
Communication	Communicate ADS intentions, status, and misbehavior to Safety Driver. Communicate with operations center.
Monitoring	Monitor the operation of the ADS via HMI internal display. Monitor HSI for missed detections, false detections, unsuitable motion plans, and poor data quality. Warn safety driver to disengage ADS. Record notes about system and road conditions and incidents for post-shift debrief.

Fleet Operations Center (FOC)

The fleet operations center is a physical space in which hired operators monitor the HD-AV fleet in a control room environment. Each fleet operator may be tasked to monitor multiple HD-AV systems through a dashboard and provide warnings to safety drivers and/or safety operators on board. In addition, fleet operators receive incident notifications automatically and play a role in traffic, route, and accident management. A description of the high-level tasks of the operators at the fleet operations center is shown in Table 7.

Table 7: Fleet Operations Center task categories and descriptions.

Task Category	**Description**
Backup Supervision	Perform a backup support role for Safety Driver and Safety Operator. Monitor fleet of vehicles and their statuses from control room environment via dashboard.
Reporting	Communicate with Safety Driver/Safety Operator about potential obstacles and risks coming ahead. Respond to alerts of Safety Driver inattention. Warn safety driver to disengage ADS. Respond to accident scenarios, alert emergency services, and participate in post-accident debrief.

A summary of the overall HD-AV high-level tasks and their divisions are compared in Table 10. As shown, the ADS and the driver divide the DDT and related tasks and coordinate this task allocation based on control transitions (Camila Correa-Jullian et al., 2024). The high-level tasks are described in Table 8.

Table 8: HD-AV high-level task descriptions.

High-Level Task	Description
Monitor DDTs	Perform Object and Event Detection Tasks
Plan DDTs	Perform the planning stage of the Object and Event Reaction tasks
Execute DDTs	Perform the execution stage of the Object and Event Reaction.
Control Vehicle	Physically control the vehicle.
Monitor Teammate	Monitor the state of the driver/operator.
Request Support	Request support from other agents to perform shared tasks.
Initiate Control Transitions	Initiate a handover or takeover request.

Four levels of engagement, described in Table 9, are used to categorize the participation of the agents in each of the high-level tasks.

Table 9: Level of engagement for HD-AV high-level tasks.

Level of Engagement	Description
Always	This task is continuously performed or available during the system's operation
Partial	This task is performed temporarily (while the ADS or manual driving is engaged) or up to a partial degree (the remote operator is engaged with multiple vehicles)
Backup	This is a safety-backup task only performed if another agent has failed to perform a main or temporary task
Never	The agent does not perform this task during operation.

Table 10: HD-AV task division summary.

Tasks/Agent	ADS	Driver	FOC
Monitor DDTs	Always	Partial	Backup
Plan DDTs	Partial	Partial	Never
Execute DDTs	Partial	Partial	Never
Control Vehicle	Partial	Partial	Never
Monitor Teammate	Always	Always	Partial
Request Support	Partial	Backup	Backup
Initiate Control Transitions	Always	Always	Never

Operational Profile

The operation of the reference fleet and the functions each of the human and machine agents perform may vary depending on certain conditions. For this work, the high-level tasks of the agents described in the previous section are organized into an operational phase diagram depicted in Figure 5. A brief description of the five phases is presented specifically for middle mile HD-AV operations. The operational profile and further modeling stages do not include the safety operator since it is not present in most planned commercial HD-AV operations.

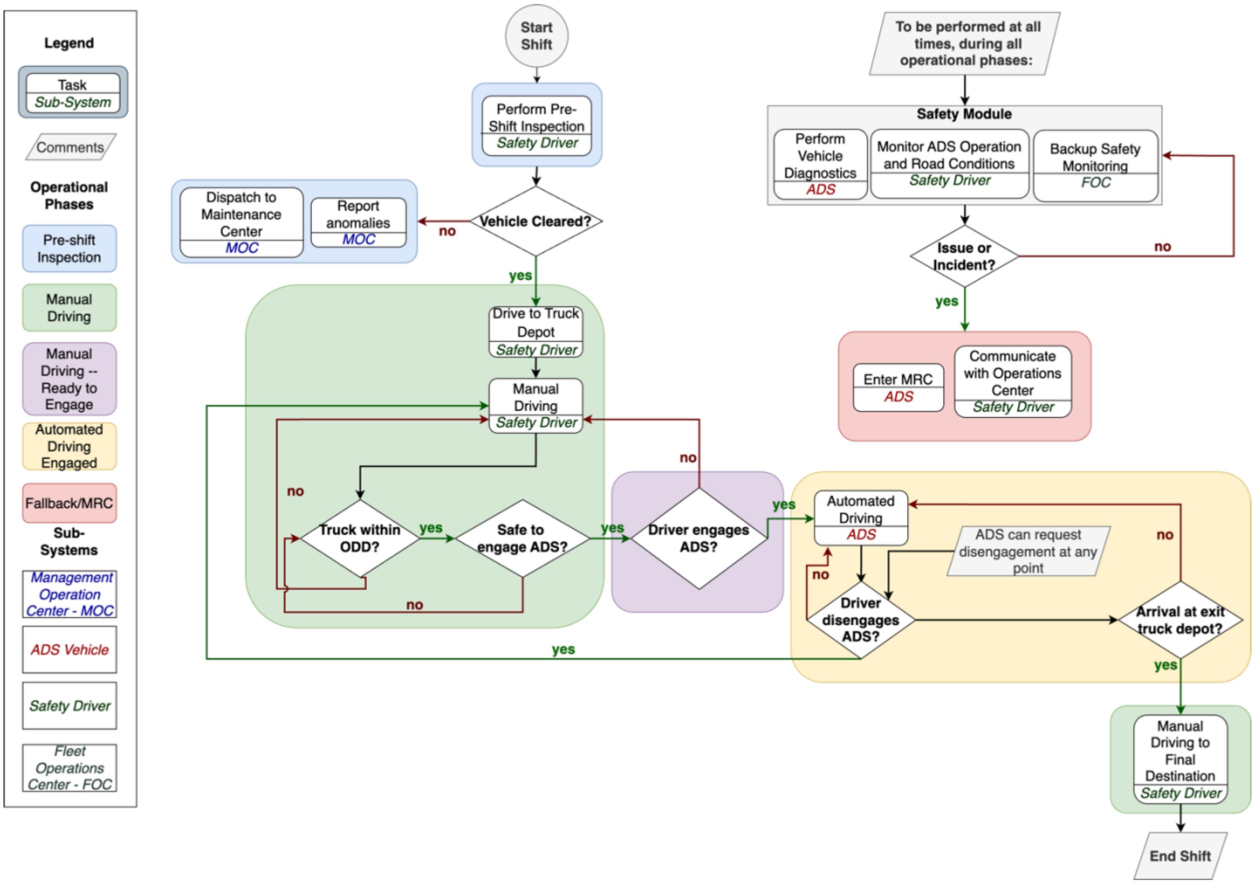

Figure 5: Diagram of operational phases and subsystems.

Pre-Shift Inspection

The first operational stage of a HD-AV is a pre-shift inspection, which is conducted by the safety driver prior to beginning the driving shift. This pre-shift inspection is typical for commercial trucks and involves inspecting safety-critical components (i.e. brakes, fluid levels, and tire pressure) and tracking this analysis through a standardized checklist. With the introduction of the ADS, this pre-shift inspection also includes a thorough inspection of the retrofitted ADS sensors, connectors, and mounts. Additionally, the safety driver is required to conduct a similar inspection every time the truck stops, for instance: during refueling, mealtimes, and breaks. If the vehicle is not in adequate condition, it is reported to the fleet operations center and the vehicle is dispatched to the maintenance operation center (MOC), whose job it is to conduct any repairs or tests that will render the vehicle operational. If the vehicle is in adequate condition, it is then approved to move on to begin the shift and proceed to Manual Driving.

Manual Driving

The first section of the driving stage is Manual Driving. The Manual Driving phase refers to the stage when the safety driver takes full control of the DDT while it is outside the ODD. This stage can involve navigating non-highway roads leading to the truck depot, or any stage during the middle-mile journey in which the truck is not in its ODD and the safety driver is in control of the vehicle. The fleet operations center monitors the vehicle’s location, status, and any upcoming obstacles on the road. This stage can be interrupted by any incidents outside expected operations, such as system failures or unplanned route changes. In this stage, the safety driver is responsible for fallbacks and incident management. If an incident occurs, this stage is interrupted, and the fleet operations center is notified. The driver monitoring system (DMS), which is a component of the ADS, monitors the drivers to ensure that they are fully aware and attentive to the driving task at hand.

Manual Driving—Ready to Engage

Manual Driving—Ready to Engage denotes when the vehicle has entered its ODD and the ADS has determined that it is feasible to activate the automated driving phase, but the safety driver remains in control. This stage is communicated to the safety driver through a combination of visual and auditory messages. In this stage, the safety driver decides whether to engage the Automated Driving. This stage can also occur if the safety driver disengages Automated Driving for any reason while the vehicle remains in its ODD. The driver has the option to reengage it while still in the Manual Driving—Ready to Engage phase. The fleet operations center continues to monitor the vehicle’s location, status, and upcoming obstacles. Like the Manual Driving phase, the safety driver is responsible for fallbacks and incident management. If an incident occurs, this stage is interrupted, and the fleet operations center is notified.

Automated Driving Engaged

This stage occurs when the ADS is performing the DDT and is controlling the steering, brake actuation, and road monitoring. The vehicle enters this stage when the safety driver activates the automated driving function via a button built into the steering wheel. In this stage, the safety driver remains in the car and monitors road conditions for any unexpected scenarios. While the vehicle is in this stage, the ADS is responsible for fallback, but the safety driver can also intervene and take manual control at this stage. For instance, this can occur when the triggers for fallback are not autonomously detected. The ADS at this stage also detects when the vehicle is nearing the limits of its ODD and alerts the driver with visual and auditory cues in this case, notifying them to resume manual control. Additionally, if the ADS detects that the driver is inattentive, the driver and fleet operations center are notified, and they can proceed accordingly. If the vehicle exits its ODD and the safety driver has not taken control, or if an incident occurs, then the vehicle enters the Fallback/MRC stage of operations.

Fallback/MRC

The fallback stage can be triggered by several incidents, for instance: internal system failures, breaches of the ODD environment, rapid changes in weather or road conditions, or incidents with other vehicles or pedestrians on the road. In the event a DDT fallback is triggered, it is expected that the ADS plans and implements a DDT fallback strategy, and achieves an MRC – i.e., the vehicle implements a safe stop, unless the safety driver intervenes to resume Manual Driving. After an MRC occurs, the post-incident procedures are triggered and the safety driver and fleet operator decide the course of action to take, whether that be remotely assisting the vehicle, or recovering it for return to the maintenance operations center. In addition, a post-shift debrief is conducted to collect information about the fallback trigger and the result of the MRC.

Scenario Modeling through ESDs

The third step of the hazard identification methodology is developing ESDs representing each operational phase and the transitions. Here the primary focus is developing qualitative ESDs representing the transitions between the Manual Driving, Ready to Engage, and Automated Driving operational phases. The critical scenarios modelled were the safety driver requesting an ADS handover (ESD 1.1), the safety driver performing a take-over (ESD 1.2), the manual driving phase transitioning to the ready to engage phase (ESD 2.1), and the ADS vehicle approaching the ODD limits (ESD 2.2). ESDs 1.1 and 1.2, representing the control transition events, were adapted from Correa-Jullian et al., 2024, and are provided in the appendix. If an end state serves as an initiating event for a different ESD, these are noted as transition states. ESDs 2.1–2.2 contain transition states to ESDs 1.1–1.2 depending on if a takeover or handover is requested. The defined end states in the ESDs, shown in Table 11, denote a successful trip (ES1), a delayed trip (EF2), or a collision risk (EF3). With real-world and simulation data, quantitative probabilities and risk levels could be attached to the events and end states in the ESD.

Table 11: End states for ESDs.

End State	**Name**	**Description**
ES1	Trip completed successfully	The HD-AV safely arrives at the destination driven by the ADS agent or on-board driver.
EF2	Trip is delayed	The trip has been interrupted by a failed takeover or handover event, leading to the driver implementing MRC or the HD-AV implementing FMS. This can lead to an interruption or delay in the shift.
EF3	Collision risk	The ADS agent or the safety driver fails to correctly perform a DDT fallback, and the HD-AV is at risk of collision. This can refer to collision with property, another vehicle, or a pedestrian.

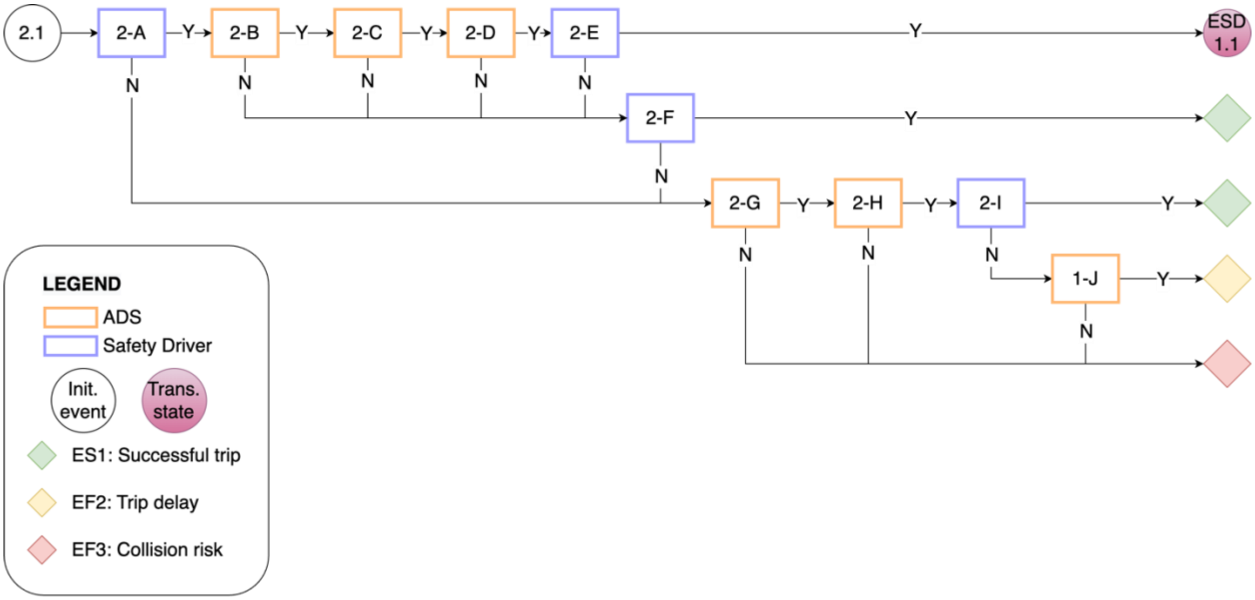

Figure 6: ESD 2.1 – Manual driving.

Table 12: ESD 2.1 – Event descriptions.

Event	**Name**	**Agent**
2.1	Vehicle cleared to begin shift.	-
2-A	Driver correctly performs manual driving task outside of ODD.	DRI
2-B	ADS correctly detects ODD entrance.	ADS
2-C	ADS determines that it is safe to engage autonomy.	ADS
2-D	ADS notifies driver of handover request.	ADS
2-E	Driver requests handover	DRI
2-F	Driver correctly performs manual driving task inside ODD	DRI
2-G	DMS detects driver inattention	ADS
2-H	DMS alerts driver of inattention.	ADS
2-I	Driver responds to alert and is able to complete DDT.	DRI
1-J	Vehicle can implement FMS.	ADS
ESD 1.1	Driver requests handover.	DRI

ESD 2.1, pictured in Figure 6, with events described in Table 12 initiates with the ADS vehicle being cleared to begin its shift. The safety driver is expected to correctly perform the manual driving task prior to the vehicle entering the ODD. If the vehicle enters the prescribed ODD, the ADS detects this entrance and whether it is safe to engage the ADS. If it is deemed safe (i.e. there is no imminent collision risk and the vehicle remains in the ODD), the ADS notifies the driver of a handover request. The driver can decide whether to approve the handover request, which then leads into ESD 1.1. If the driver decides to not approve the handover request, they are responsible for completing the DDT, even with the vehicle in the ODD. This can also lead to a successful trip although automated driving was not engaged. If the safety driver does not correctly perform the DDT, whether this occurs inside or outside of the ODD, the DMS, a component of the ADS, nominally detects and alerts the driver of inattention. If the driver responds to the DMS alert and can complete the driving task, then this is denoted as a successful trip. Alternately, the vehicle can implement FMS in the case of driver inattention, leading to a trip delay. If the vehicle is unable to successfully implement FMS, this can lead to a vehicle collision risk.

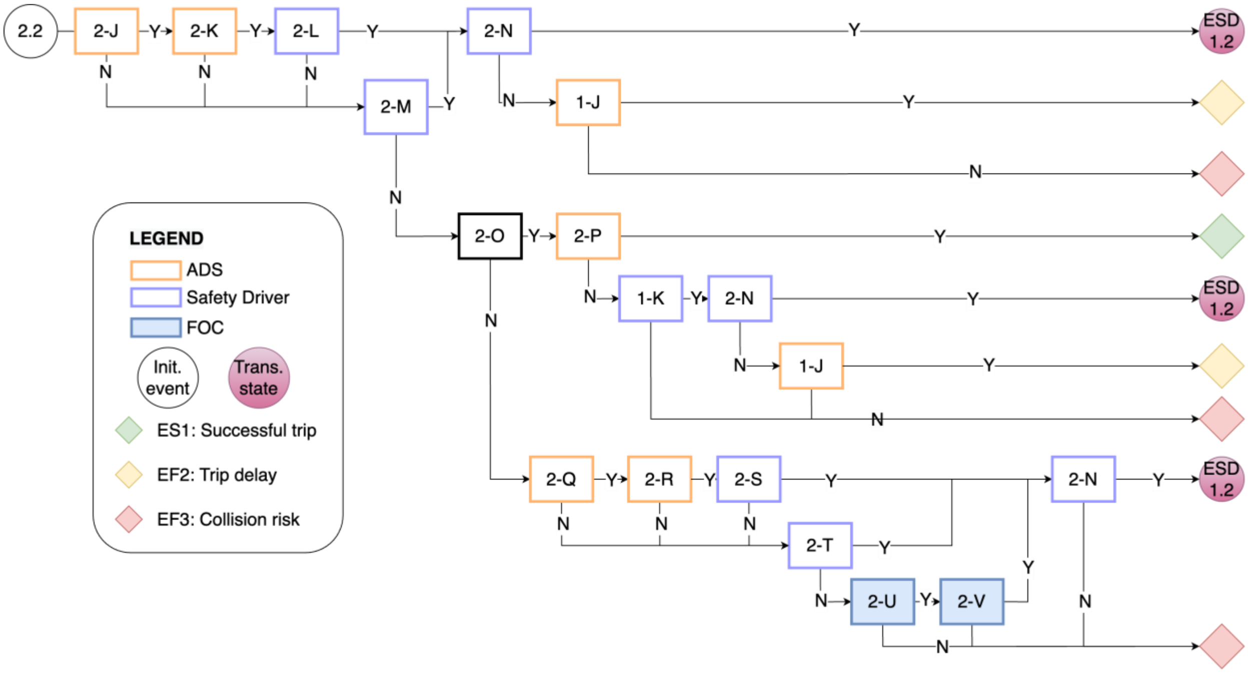

Figure 7: ESD 2.2 – Vehicle approaches limits of ODD.

Table 13: ESD 2.2 – Event descriptions.

Event	Name	Agent
2.2	Vehicle approaches limits of ODD.	-
2-J	ADS detects ODD limit approach.	ADS
2-K	ADS notifies driver that vehicle is approaching ODD limits and requests takeover.	ADS
2-L	Driver detects ODD limit approach and ADS takeover request.	DRI
2-M	Driver detects ODD limit approach.	DRI
2-N	Driver performs takeover of vehicle.	DRI
1-J	Vehicle can implement FMS.	ADS
2-O	Vehicle remains in ODD.	-
2-P	ADS is able to perform the entire DDT within ODD.	ADS
1-K	Driver detects that DDT-fallback is needed.	DRI
2-Q	ADS detects ODD limit exit.	ADS
2-R	ADS notifies driver and FOC that vehicle has exited ODD and requests takeover.	ADS
2-S	Driver detects notification of ODD exit and ADS takeover request.	DRI
2-T	Driver detects ODD limit exit.	DRI
2-U	FOC detects ODD limit exit.	FOC
2-V	FOC notifies driver that vehicle has exited ODD.	FOC
ESD 1.2	Driver performs take-over.	DRI

The event sequence diagram pictured in Figure 7 and described in Table 13 starts with automated driving engaged, and the ADS vehicle approaching the limits of the ODD. The ADS is expected to detect the ODD limit approach and notify the safety driver. The driver may also preventatively respond to DDT fallback triggers, independent of whether the ADS alerts the driver. The safety driver could choose to take over, leading to ESD 1.2. If the driver does not detect the limit approach, the ADS is expected to implement a failure mitigation strategy (FMS), which involves allowing the vehicle to come to a safe stop, leading to a trip delay. Hence, the safety barriers are constructed hierarchically, with the first being ADS detection, then driver detection and intervention, and finally implementing FMS. If the vehicle exits the ODD, the ADS is expected to alert the safety driver. Even if the ADS does not detect or alert the driver of the ODD breach, the safety driver may independently detect and respond to the ODD breach. If neither the safety driver nor the ADS detect the ODD breach, and the FOC receives a breach notification, the FOC can notify the driver of the need to perform a takeover. If the FOC does not detect the ODD breach, there is an issue in performing the takeover, or FMS fails, this may lead to an unmitigated risk of collision.

Modeling critical tasks through CoTA

The fourth step involves modelling agents' tasks and interactions through CoTA. CoTA is a technique which decomposes high-level goals of each agent into a series of tasks and subtasks to analyze interactions between agents and how they relate to overall goal success. It decomposes tasks based on Hierarchical Task Analysis (HTA), where tasks are re-described until fundamental tasks relating to the interactions between agents appear (Ramos et al., 2020b). These tasks are categorized based on the IDA (Information, Decision, Action) cognitive model (Chang & Mosleh, 2007). IDA was initially developed for human agents, where tasks are decomposed into receiving information from a system (I), planning a course of action (D), and performing the action (A), but it has been also applied to the maritime industry in the context of Autonomous Ships (AS). Each task and subtask in CoTA is categorized into a series of plans which denote the order in which the tasks are carried out. The task categories comprising these plans are summarized in Table 14.

Table 14: CoTA task plan categories.

Task Type	Symbol	Meaning
Parallel	1 // 2	Perform Tasks 1 and 2 at all times
Sequential	1 -> 2	Perform Task 1 then Task 2
Triggered	1 -- 2	Perform Task 2 only if triggered by Task 1.
Exclusive	1 or 2	Perform Task 1 or Task 2.

CoTA models were developed for the Safety Driver, ADS, and FOC agents in this system. For illustration purposes, simplified versions of the Safety Driver (DRI) and ADS CoTAs are presented in Figures 8-9, but as needed, connections to other agent CoTAs are shown. The complete FOC CoTA is shown in Figure 10. Relevant CoTA subtasks used for metric creation are also shown in the appendix.

The role of the safety driver in the HD-AV was decomposed into six high-level tasks described in Table 15. These tasks involve continuous monitoring of driving conditions (Task 1) and communicating with the FOC (Task 6) when required. During the Manual Driving operational phase, the safety driver is responsible for performing DDT planning and execution (Task 2). Even while the ADS is engaged, the safety driver is expected to use the information from Task 1 to determine whether a DDT fallback is needed (Task 3). If a DDT fallback is needed, Task 3 triggers Task 4, executing the DDT fallback plan. Throughout all operational phases, the safety driver also interacts with the ADS vehicle (Task 5). The driver monitors the control status, can request transitions and receives transition requests from the ADS, and transmit and receive ADS alerts. The safety driver’s Task 5 interfaces with ADS’s Task 5, where the ADS also monitors control status, can request control transitions, and receive handover requests, and transmits and receives driver alerts.

Table 15: Safety Driver CoTA tasks.

Num.	Subtask	Type	Description
1	Monitor driving conditions	Parallel	The safety driver performs monitoring tasks during all phases of operation. This involves monitoring the ADS vehicle operation, driving environment, alerts from the ADS vehicle, and communications from the FOC. Information gathered from this task supports the other tasks.
2	Perform DDT planning and execution	Triggered	This occurs during the Manual Driving phase or is triggered by Task 5, when a control transition occurs to transfer DDT control to the safety driver. When triggered, the driver uses the information from task 1 to fully plan and execute the DDT. The DDT involves employing OEDR functions, following local traffic rules, and, if necessary, implementing tactical maneuvers.
3	Determine if a DDT fallback is required	Parallel/ Trigger	At all operational phases, the safety driver determines if the situation requires a DDT fallback plan. A DDT fallback plan can be triggered by an ODD breach or limit approach, a vehicle or sensor failure, or by a perceived risk of collision.
4	Execute DDT fallback plan	Sequential/Triggered	This task is triggered by Task 3. The driver determines the DDT fallback strategy and implements a fallback plan. The strategy requires the driver to assess the vehicle condition and determine what the end state should be, either allowing the ADS to continue performing the DDT, requesting a control transition, or requesting an emergency stop. Once planned, the driver implements the fallback plan and evaluates the outcome.
5	Interact with the ADS vehicle	Parallel	The driver continuously receives and transmits commands to the ADS regarding vehicle control transitions, emergency stop requests, and navigational inputs. Additionally, here the driver manages the vehicle control transitions. The driver may request control transitions, i.e., driver-initiated handovers or takeovers and respond to system-initiated requests.
6	Communicate with FOC	Parallel	At all operational phases, the safety driver communicates with the fleet operations center. For this task, the safety driver receives communications from the FOC, plans a response, and then responds to the FOC.

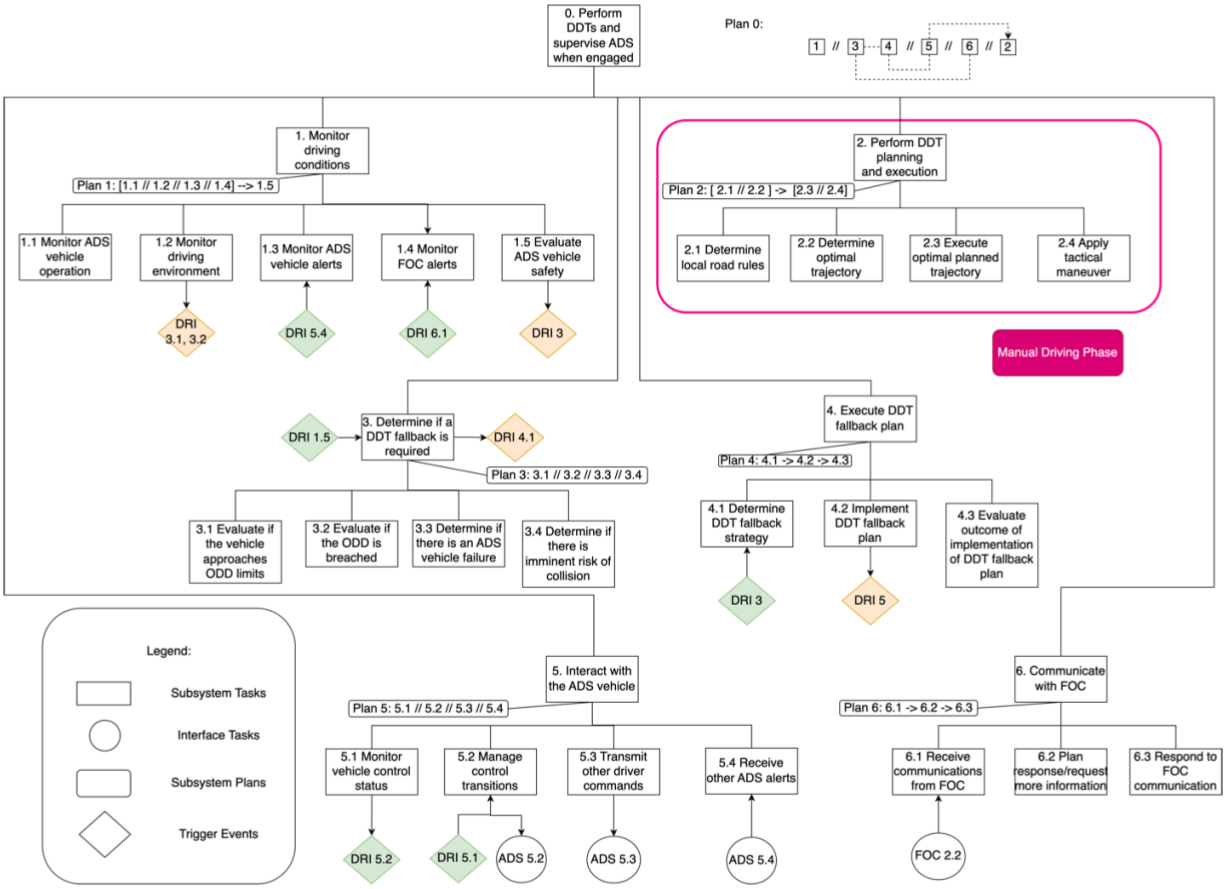

Figure 8: Simplified Safety Driver CoTA model.

The role of the ADS in HD-AV operations was decomposed into seven high-level tasks described in Table 16. These tasks involve continuous analyzing sensor data to observe the driving environment (Task 1) and performing self-diagnostic tasks to observe ADS operation (Task 6). During the Automated Driving operational phase, the ADS is responsible for performing DDT planning and execution (Task 2), determine whether a DDT fallback is needed (Task 3), and if needed, executing the DDT fallback plan (Task 4). Throughout all operational phases, the ADS also interacts with the safety driver (Task 5). The ADS transmits and receives alerts to the driver regarding control transition events, emergency stops, and navigational inputs, directly interfacing with Task 5 of the safety driver. Additionally, the ADS communicates with the safety driver and FOC about ADS and physical vehicle status, also initiating DMS alerts (Task 7).

Table 16: ADS CoTA tasks.

Num.	Subtask	Type	Description
1	Perform DDT OEDR supporting functions	Parallel	The ADS gathers and processes sensor data to gain information about the vehicle, environment, and ODD to support the other parallel tasks.
2	Perform DDT planning and execution	Triggered	This task is triggered by Task 5, when a control transition occurs to transfer DDT control to the ADS. The ADS continuously uses information from Task 1 to fully plan and execute the DDT.
3	Determine if a DDT fallback is required	Parallel/ Trigger	At all operational phases, the ADS continuously determines if the situation requires a DDT fallback plan, which can be triggered by an ODD breach or limit approach, a vehicle or sensor failure, a perceived risk of collision, or by an emergency stop request initiated by the driver.
4	Execute DDT fallback plan	Sequential/Triggered	This task is triggered by Task 3. The ADS determines the DDT fallback strategy and implements a fallback plan. The strategy can involve continuing the DDT, implementing MRC or SSC. Once planned, the ADS implements the fallback plan and evaluates the outcome.
5	Interact with safety driver	Parallel	The ADS continuously receives and transmits commands to the driver regarding vehicle control transitions, emergency stop requests, and navigational inputs.
6	Perform self-diagnostic tasks	Parallel	The ADS monitors its subsystems and sensor data to determine if there are any malfunctions in the hardware or software. To do this, it performs self-diagnostic tests that notify the driver.
7	Communicate with safety driver and FOC	Parallel	At all operational phases, the ADS alerts the safety driver and the FOC about ADS status, vehicle status, and DMS system alerts.

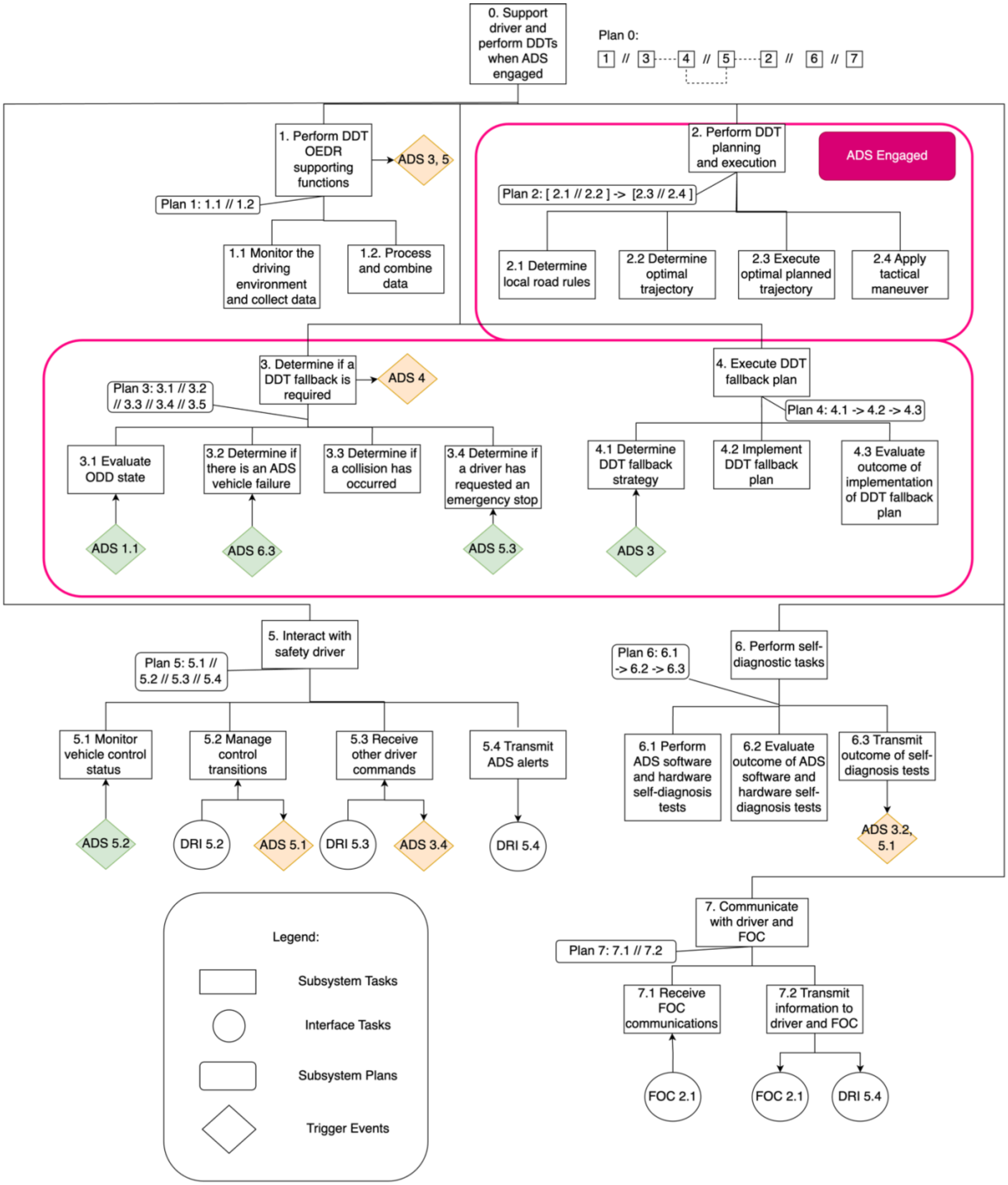

Figure 9: Simplified ADS CoTA model.

The role of the FOC was decomposed into three high-level tasks described in Table 17. These tasks involve continuous monitoring of operations (Task 1), including ADS, physical vehicle, and driver monitoring alerts. Additionally, the FOC has access to information about ODD exits so that operators are aware of the physical location and status of the HD-AVs. The next parallel task is communicating with the driver and ADS (Task 2) about status and vehicle alerts. If there are any issues outside of nominal operation, the FOC provides a backup support role and can communicate with the driver about fallback procedures, however the FOC does not provide direct vehicle input or control. Finally, at all stages the FOC assesses the need for and initiates incident response support tasks (Task 3). Again, while the FOC does not intervene with the vehicle, they coordinate with third party individuals that can provide incident support and coordinate information for a post-incident debrief.

Table 17: FOC CoTA tasks.

Num.	**Subtask**	**Type**	**Description**
1	Monitor operations	Parallel	The FOC views information about ADS, physical vehicle status, and driver status on a control room dashboard. While this information is limited due to the necessity to monitor multiple vehicles, the FOC receives the information that is relevant.
2	Communicate with driver and ADS	Parallel	At all operational phases, the FOC communicates with the ADS and the safety driver. The FOC receives DMS alerts and other communications from the ADS through the dashboard and verbal communications from the safety driver.
3	Incident response	Parallel/ Triggered	At all operational phases, the FOC continuously determines if there is a need for incident response, which can be triggered by a vehicle or sensor failure, a perceived risk of collision, an actual collision, or by an emergency stop request initiated by the driver. The FOC communicates with first responders and coordinates data collection for the debrief. Additionally, if there is a need to collect the vehicle and transport it back to the maintenance operations center, the FOC coordinates this aspect of the incident response.

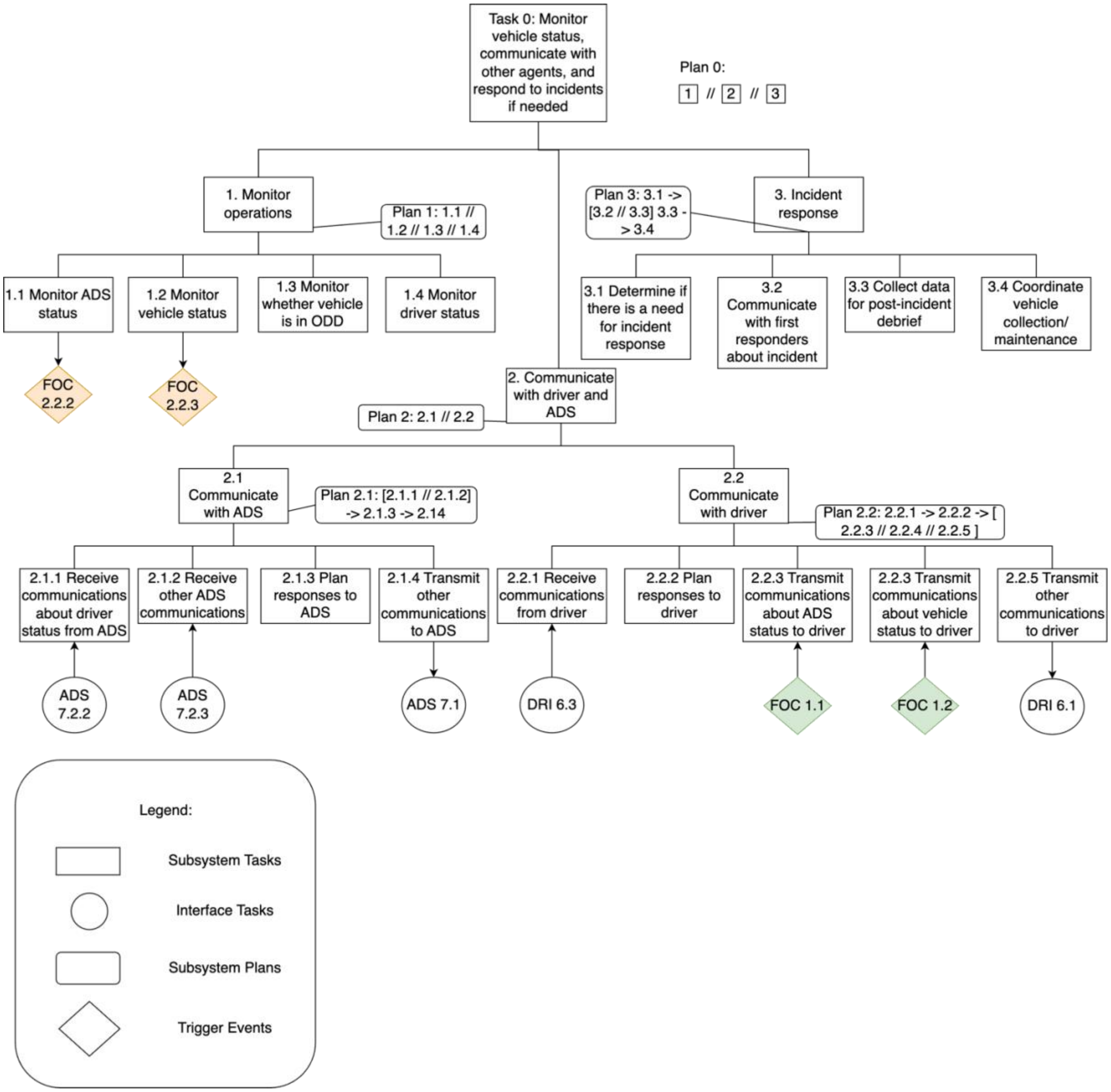

Figure 10: Complete FOC CoTA model.

As shown in Tables 15 and 16, the safety driver and ADS are responsible for similar high-level tasks, such as performing aspects of the DDT planning, execution, and fallback, but they perform these at different operational phases, dictated by control transitions. The safety driver performs DRI Task 2 only during the Manual Driving Phase, and the ADS performs ADS Task 2 only during the Automated Driving Engaged phase. Both the safety driver and ADS can assess the need for and implement a DDT fallback at any stage of operation. Task 5 for both DRI and ADS include the control transitions that lead to changes in the operational state of the subsystem. The distinction in driver and ADS DDT fallback is that driver DDT fallback can lead to a takeover request or emergency stop request, but the ADS fallback leads to a Minimal Risk Condition (MRC) or Stable Stopped Condition (SSC), which are initiated by the vehicle and can lead to an operational delay.

To demonstrate the interfacing between separate agent CoTAs for the case of takeover and handover events, the full decomposition of subtask 5.2.1 for the safety driver and ADS is shown in Figures 11 and 12. Task 5.2.1 is a subtask of Task 5.2, which denotes the “Manage control transitions” task. This task is divided into managing driver-initiated and system-initiated control transition events. For the case of driver-initiated takeovers, the driver first determines whether a takeover is needed or desired using information from DRI Task 5.1.2, which takes information from DRI Task 1 (driving conditions) and DRI Task 3 (assessing the need for a DDT fallback). If a takeover is desired, the driver performs the takeover (DRI 5.2.1.1.2), and the ADS is required to detect the takeover input (ADS 5.2.1.1.1). For the driver-initiated handover, first the driver determines whether an ADS handover is feasible (DRI 5.2.1.2.1) and initiates a request (DRI 5.2.1.2.2). This interfaces with ADS 5.2.1.2.1, which is receiving the handover request, and ADS 5.2.1.2.2, which is the ADS determining whether the handover is feasible. The ADS then communicates the status of the request, which can be an approval or a denial of the request (ADS 5.2.1.2.3). The driver is responsible for determining whether the handover is approved (DRI 5.2.1.2.3), and if it is approved the ADS performs the desired handover (ADS 5.2.1.2.4). Once the handover occurs, the ADS transmits a confirmation (ADS 5.2.1.2.5), which the driver is responsible for detecting (DRI 5.2.1.2.4). It can be noted that a takeover or handover can be unsuccessful due to a lack of detection, an ADS operation failure, or other issue, and unsuccessful rates of takeovers and handovers are quantified. The control transition can be attempted again after the fact, or if there is a critical issue, the Fallback/MRC stage or an emergency stop can be triggered. The CoTA provides a structured way to observe agent interactions on the same hierarchy and assess which of these interactions can point to operational risks.

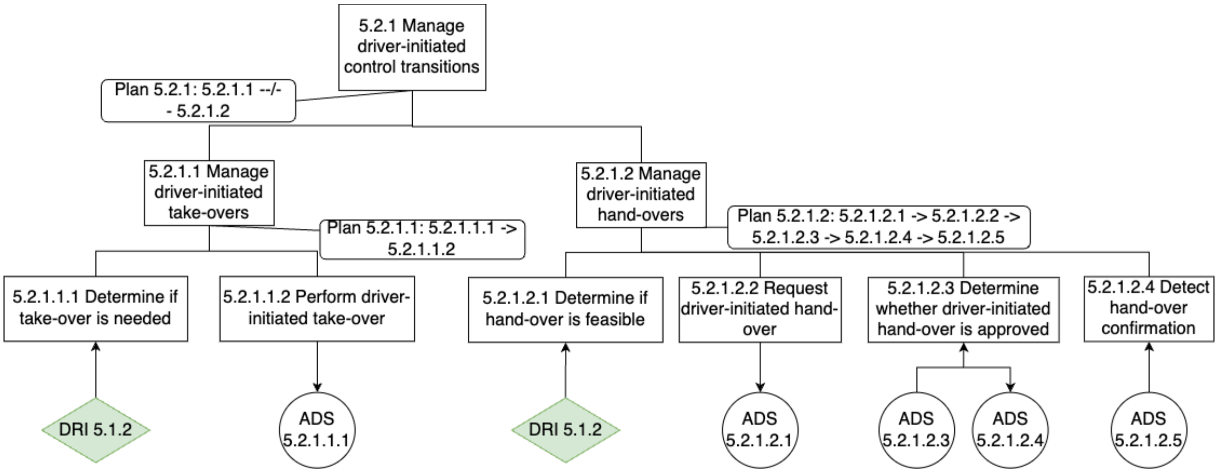

Figure 11: DRI Task 5.2.1 CoTA decomposition.

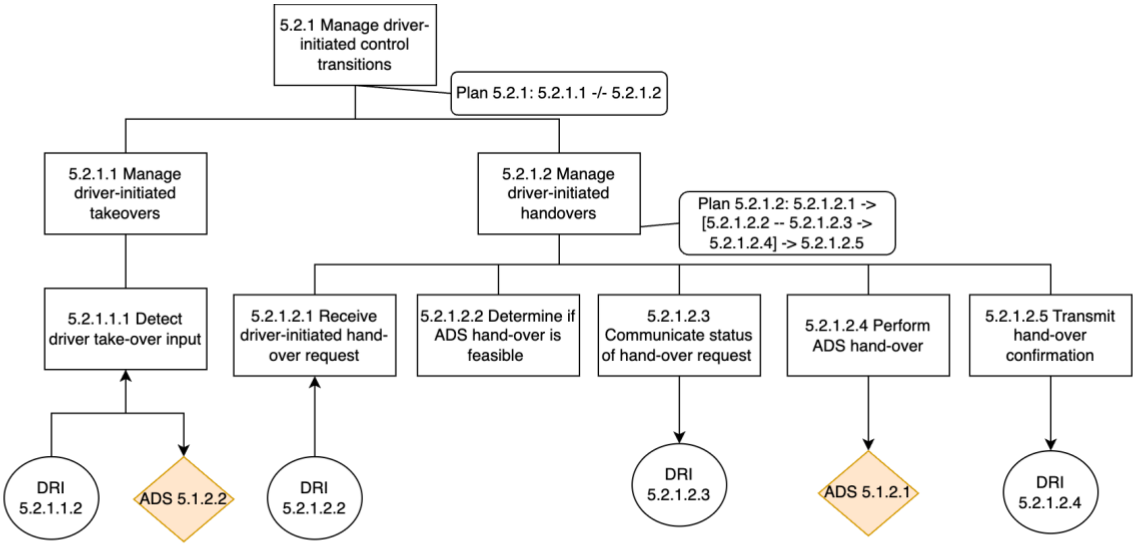

Figure 12: ADS Task 5.2.1 CoTA decomposition.

Table 18: Interface tasks and triggering Events for DRI/ADS CoTA 5.2.1.

Agent	Task Num.	Task
DRI	5.1.2	Determine if vehicle control change is desired
ADS	5.2.1.2.2	Determine if driver is in control of the vehicle
ADS	5.2.1.2.1	Determine if ADS is in control of the vehicle

Derived Human-System Interaction Metrics

In observing failure paths from the ESD and the tasks that leading to failure modes in the CoTA, a list of human-system interaction metrics was developed and grouped by categories. The metrics were created by determining which factors of the CoTA tasks could be measured to point out potential operational safety weaknesses. The proposed metrics for the modelled HD-AV system are summarized in Tables 19-23. If not included in earlier figures, the subtasks with the CoTA origins for these metrics are in the appendix. Ideally, these metrics can be tracked in HD-AV simulation and testing to inform design of components like HMI and operational tasks in initial stages and validate the system in later testing stages. Additionally, many metrics can be used during nominal operation and are not incident-based, so data trends can be assessed. Most of these metrics would be collected from the ADS data log, which keeps track of sensor and alarm data, operational phase changes including takeover and handover events and quality of takeovers based on vehicle dynamics. For qualitative metrics, questionnaires, surveys, and interviews can be employed to determine possible root causes leading to decisions made by the safety drivers and FOC operators assess the quality of the ADS post-incident response.

Table 19: Control transition metrics.

#	Name	Definition	Unit	Data source	Use case	Origin
1	Rate of Successful Driver-Initiated Handovers	Ratio of successful driver handovers to total number of handover requests.	%	Data log	Testing/ operation	DRI CoTA 5.2.1.2.4, ADS CoTA 5.2.1.2.5, ESD 1.1
2	Rate of Successful Driver-Initiated Takeovers	Ratio of successful driver takeovers to total number of takeover attempts.	%	Data log	Testing/ operation	DRI CoTA 5.2.1.1.2, ADS CoTA 5.2.1.1.1, ESD 1.2
3	Rate of Successful System-Initiated Handovers	Ratio of successful driver handovers to total number of system-initiated handover requests.	%	Data log	Testing/ operation	DRI CoTA 5.2.2.2.3, ADS CoTA 5.2.2.2.5
4	Rate of Successful System-Initiated Takeovers	Ratio of successful driver takeovers to total number of system-initiated takeover requests.	%	Data log	Testing/ operation	DRI CoTA 5.2.2.1.2, ADS CoTA 5.2.2.1.3
5	Rate of ADS Handover Approval	Ratio of driver approval of system-initiated handovers to system-initiated handover requests.	%	Data log	Testing/ operation	DRI CoTA 5.2.2.2.2
6	Reason for Driver-Initiated Takeover	Category for reason safety driver initiated a takeover (e.g. lack of trust, unnoticed ODD breach).	n/a	Survey or interview	Testing/ operation	ESD 1.2, DRI CoTA 5.2.1.1
7	Reason for System-Initiated Takeover	Category for reason ADS initiated a takeover request (e.g. ODD breach, collision risk).	n/a	Data log	Testing/ operation	ADS CoTA 5.2.2.1.1
8	Quality of Takeover (TTC-Based)	Minimum time to collision and maximum resulting lateral and longitudinal acceleration after the initiated takeover request.	sec	Data log	Testing/ operation	Literature/ system model
9	Quality of Takeover (Dynamics-Based)	Maximum resulting weighted sum of lateral and longitudinal acceleration after the initiated takeover request.	m/s^2	Data log	Testing/ operation	Literature/ system model
10	Quality of Takeover (TOT-Based)	Takeover time (TOT) interval between takeover request (TOR) and the driver’s first maneuver	sec	Data log	Testing/ operation	Literature/ system model

Within the control transitions category, the metric “Rate of Successful Driver-Initiated Takeovers” was determined by observing the interaction between DRI and ADS Task 5.2.1.1, shown in Figures 11 and 12. A failure in detection of driver takeover input would lead to an unsuccessful takeover, which could highlight a potential software or hardware ADS risk. Hence, recording the possible root causes of the driver-initiated takeover can then support system design improvement decisions (e.g., control transition mechanisms) or temporary restrictions in the ODD during operation.

Further control transition metrics relate to success rates for other control transition events, and the reasons these control transitions occur. Although the reason for a control transition can be obtained qualitatively through surveys and interviews, by gaining information from initial studies, the potential reasons can be categorized and then tracked in a more systemic manner. In terms of quality of takeover measurements, there has been research into the creation of standardized metrics, three of which are chosen here (Cao et al., 2021).

Table 20: Alert-related metrics.

#	Name	Definition	Unit	Data source	Use case	Origin
1	Alerts Resulting from Vehicle-Related Malfunctions	Ratio of alerts coming from vehicle sensor, ADS or vehicle malfunction to total number of alerts	%	Data log	Testing/ operation	DRI CoTA 5.4.2
2	Alerts Resulting from Onboard Safety Driver	Ratio of alerts generated by the driver monitoring system to total number of alerts	%	Data log	Testing/ operation	DRI CoTA 5.4.3
3	Alerts Resulting from Environment	Ratio of alerts coming from road conditions, ODD breach to total number of alerts	%	Data log	Testing/ operation	DRI CoTA 5.4.1
4	Alerts Not Acted On	Ratio of alerts not responded to by safety driver to total number of alerts	%	Data log	Testing/ operation	ESD 2.2, Event 2-I

The alert-related metrics were used to measure the ratio of alerts arising from diverse sources (i.e. vehicle-related malfunctions, driver monitoring system, ODD breaches, etc.) as well as the ratio of alerts not responded to by the safety driver. For instance, while it is important to assess whether alerts are detected by the safety driver, this presents significant difficulties. Hence, the metric “Alerts Not Acted On” can serve as a partial indicator that indicate a need for developing improvements being at HMI display, alert design, or driver training level.

Table 21: Incident-related metrics.

#	Name	Definition	Unit	Data source	Use case	Origin
1	Incident Rate - Driver in Control	Rate of incidents (collision, property damage) leading to an initiation of post-incident procedures while safety driver is in control	#/vmt	Data log	Operation	FOC CoTA 3
2	Incident Rate - ADS in Control	Rate of incidents (collision, property damage) leading to an initiation of post-incident procedures while ADS is in control	#/vmt	Data log	Operation	FOC CoTA 3
3	Incident Rate – Driver in Control – Disruption Only	Rate of incidents while safety driver is in control that only result in a traffic disruption	#/vmt	Data log	Operation	FOC CoTA 3
4	Incident Rate – Driver in Control – Property Damage Only	Rate of incidents while safety driver is in control that only result in property damage	#/vmt	Data log	Operation	FOC CoTA 3
5	Incident Rate – Driver in Control – Collision	Rate of incidents while safety driver is in control that result in a collision with a pedestrian or vehicle	#/vmt	Data log	Operation	FOC CoTA 3
6	Incident Rate – ADS in Control – Disruption Only	Rate of incidents while ADS is in control that only result in a traffic disruption	#/vmt	Data log	Operation	FOC CoTA 3
7	Incident Rate – ADS in Control – Property Damage Only	Rate of incidents while ADS is in control that only result in property damage	#/vmt	Data log	Operation	FOC CoTA 3
8	Incident Rate – ADS in Control – Collision	Rate of incidents while ADS is in control that result in a collision with a pedestrian or vehicle	#/vmt	Data log	Operation	FOC CoTA 3

The incident metrics track incident rate per vehicle miles traveled (vmt), and these rates are classified into incidents with the driver in control and ADS in control. Additionally, they are disaggregated into incident severity levels, with the levels being “Traffic Disruption Only”, "Property Damage Only", and "Collision", which includes incidents with damage to other vehicles, fatalities, and injuries. These are all derived from FOC CoTA Task 3, which denotes incident response.

Table 22: Fallback-related metrics.

#	Name	Definition	Unit	Data source	Use case	Origin
1	Rate of Emergency Stop Requests	Rate of times a driver requests an emergency stop (per vehicle miles driven).	#/vmt	Data log	Testing/ operation	DRI CoTA 5.3.2
2	Rate of ADS-Initiated FMS	Rate of ADS initiated FMS being implemented per mile	#/vmt	Data log	Testing/ operation	ADS CoTA 4
3	Rate of Safety Driver Manual Fallback	Rate of safety driver takeovers when fallback is needed	#/vmt	Data log	Operation	DRI CoTA 4
4	ADS Fallback Due to Driver Inaction	Rate of ADS initiated fallback, specifically when driver monitoring system (DMS) detects inattention and lack of response	#/vmt	Data log	Operation	ADS CoTA 4
5	Delay Resulting from Fallback	Per-shift time delay resulting from implementation of fallback procedures	hours	Data log	Testing/operation	FOC CoTA 3

The fallback metrics refer to rates of FMS, emergency stops, and cases of ADS fallback resulting from driver inaction. Additionally, the "Delay Resulting from Fallback" aims to quantify the physical time delay resulting from fallback procedures. The fallback metrics can be used to assess fallback scenarios, even though an incident may not occur and determine the rates at which the ADS and the safety driver initiate fallback procedures.

Table 23: Human-ADS trust metrics.

#	Name	Definition	Unit	Data source	Use case	Origin
1	Divergence from ADS Decisions	Ratio of maneuvers conducted by safety driver during Automated Driving phase differing from ADS maneuver intention	%	Data log	Testing	DRI CoTA 5.1.2
2	Time Spent in Ready to Engage Phase	Time spent in the Manual Driving—Ready to Engage phase without engaging ADS	hours/vmt	Data log	Testing/ operation	System model

The Human-ADS Trust metrics consist of measurements of disagreement with ADS maneuvers and a measure of the time spent in the Ready to Engage phase without engaging the ADS. These trust metrics can also be supplemented with existing human factors studies on ADS trust (Yang et al., 2018).

While the derived human-system interaction metrics serve as a starting point for observing trends in HD-AV system behavior, they may not be relevant for all HD-AV operational scenarios. Thus, the way they are derived from the ESD scenario modelling and CoTA task decomposition can be utilized to develop metrics for a more specific HD-AV configuration. The aim in creating these metrics is to implement leading measures that can point to potential hazards, as well as increase a focus on operational safety and trends rather than solely focusing on functional safety. While many metrics have been developed and validated specifically for ADS systems as well as for certain control transition events, there is a lack of implementation of human-system interaction metrics in a systemic way. Arriving at these measures of human-system interaction using task and agent-based models allows for a direct observation of the interactions that occur and how they can affect the overall system safety.

Discussion

Although not all events leading to potential risk can be directly measured, safety metrics can point to contributing factors that can be addressed to manage the system’s risk. Further, leading metrics can be measured prior to incidents occurring, providing a proactive view of risk assessment. For instance, while it is important to assess whether alerts are detected by the safety driver, this presents significant difficulties. Hence, the metric “Alerts Not Acted On” can serve as a partial indicator that indicate a need for developing improvements in HMI display, alert design, or driver training level.

One of the primary topics that has been assessed in ADS research is the degree of trust that human operators inside and outside the vehicle have on the system (Yang et al., 2018). Although somewhat of a qualitative assessment, trust can be estimated in a numerical sense by employing metrics in addition to qualitative information gleaned from surveys and interviews. Another consideration of human system interaction for heavy duty vehicles involves the team of human operators onboard, if there is both a safety driver and a safety operator. The dynamics between these two operators can affect the decision-making and action-taking processes regarding the ADS. For instance, if a safety operator notes a potential collision risk or need for takeover, but the driver ignores them and is reliant on either their own perspective or that of the ADS sensors, these disagreements may develop into more complex hazard scenarios. Although the disagreements between safety driver, safety operator, and ADS can be observed in a numerical manner denoting number of times a disagreement occurs, it is likely that a more complete picture will be formed through incorporating a post-shift debrief, in which questionnaires will assess the motivation behind such a disagreement.

Model-based approaches to metric creation can allow for analysis of low-level tasks that can point to areas for system safety improvement. A combined approach relying on model-based risk assessment and benchmarked simulators, such as CARLA or OpenCDA, can lead to overall improvements prior to the system development and implementation (Dosovitskiy et al., 2017; Xu et al., 2021, 2023). The advantage of combining simulation- and model-based approaches lies in the systematic methods available to model the system’s hardware, software, and human elements. Model-based approaches provide traceability as opposed to purely data-driven metrics derived directly from simulations or testing, with the added benefit of being able to be integrated into early design stages, and evolving during system development, certification, and operational phases. Indeed, developing quality safety metrics can play a significant role when assessing operational safety at later stages of system deployments, assessing their evolution over extended periods of time or miles driven.

The autonomy level of current commercial HD-AV systems is commonly advertised at an SAE Level 4, which is defined as the ADS performing the entire DDT and all associated fallback procedures. However, in current testing scenarios, there is always a safety driver in the vehicle who can intervene in case of emergencies. The active role of the safety driver is not expected to be removed due to current automated vehicle legislation and recent events in driverless passenger transport applications (Mickle et al., 2023). Thus, the effective level of autonomy may decrease to Level 3 according to current J3016 definitions; however, since the ADS nominally plans and implements the DDT fallbacks, further discussions may be required to assess intermediate levels of automation. There has been criticism of the currently defined SAE Levels, especially that their strictly linear progression overlooks the necessary hybridity of interactions between human and automated technologies (Hopkins & Schwanen, 2021). The HD-AV system can be referred to as existing at a nominal Level 4 of automation; namely it has characteristics of a Level 3 system, however at a nominal state it can perform all the tasks expected of a Level 4 system. As HD-AV companies more clearly define roles and tasks, the autonomy state and the presence of safety drivers and operators can be reassessed.

Since the interplay between safety drivers and the ADS plays a large role in the division of tasks and operation of heavy-duty systems, there is a need to include safety metrics about how these humans and systems interact. This can help inform the designs of ADS systems in turn; for instance, making HMIs more effective for safety drivers, developing more robust driver monitoring systems, and designing alert functions. In addition, observing these metrics can help with designing HD-AV operations for both efficiency and safety without promoting one at the expense of the other.

A focus on human-centered metrics has been applied in various industries such as nuclear and maritime, to develop effective human reliability assessments (HRAs) which are conducted through models such as Phoenix, SPAR-H, and IDAC (Chang & Mosleh, 2007; Ekanem et al., 2016). Existing HRA methods can be expanded upon to also incorporate automated vehicles in their models and potentially be used to expand current Probabilistic Risk Assessments (PRAs) conducted to inform the design and regulation of these automated driving systems. Additionally, many industries such as occupational health and safety and commercial aviation have been beginning to prioritize leading metrics as a proactive view of operational safety (Bayramova et al., 2023; Sheehan et al., 2016) A combination of leading metrics and human-system interaction metrics represents a necessarily push to operational safety for novel automated driving technologies and should be implemented in the upcoming HD-AV industry.

Conclusion

With the increased interest in incorporating ADS into heavy-duty commercial operations, the role that human-autonomy teams play has not been fully assessed in the HD-AV framework. Due to regulatory and legal framework, safety drivers will likely continue to be involved in HD-AV operations beyond testing, therefore it is necessary to observe the interactions between the human and machine agents in this system to assess their safety. Using a comprehensive set of human-system interaction metrics will inform operational and system design during preliminary testing phases. These metrics can improve the design of components such as HMI in the vehicle and inform the design of operational tasks. Additionally, metrics can assess trends or point to needed changes during road testing stages and eventual public deployment. This work demonstrates a methodology to use ESD and CoTA models to derive human-safety interaction metrics that can inform design and development of safe HD-AV systems. Future work can be done towards incorporating STPA and other methods to provide alternative characterization of the HD-AV systems, leading towards a more comprehensive hazard identification analysis and surrogate safety metric construction. The metric derivation methodology presented here can be adapted by HD-AV fleet operators to reflect the respective company's specific operational scenarios and agent tasks. Additionally, the evolution of selected metrics can be observed over time or miles driven in operation in order to make decisions and assess changes made.

Appendix

The appendix includes the control transition ESDs representing the takeover and handover events, and CoTA model diagrams relevant for the creation of the derived human-system safety metrics.

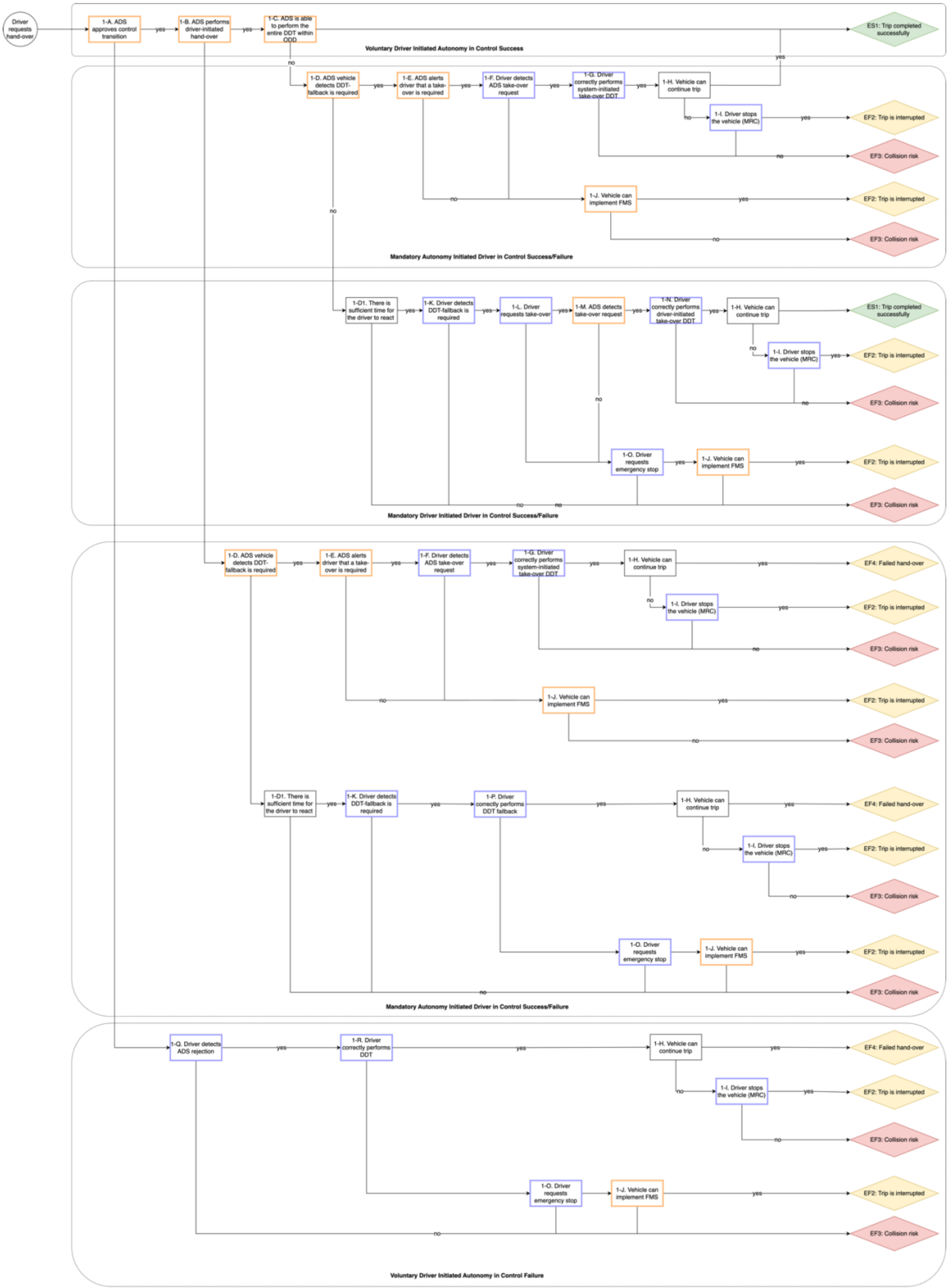
Driver requests hand-over
1-A. ADS approves control transition
1-B. ADS performs driver-initiated hand-over
1-C. ADS is able to perform the entire DDT within ODD
ES1: Trip completed successfully
Voluntary Driver Initiated Autonomy in Control Success
1-D. ADS vehicle detects DDT-fallback is required
1-E. ADS alerts driver that a take-over is required
1-F. Driver detects ADS take-over request
1-G. Driver correctly performs system-initiated take-over DDT
1-H. Vehicle can continue trip
1-I. Driver stops the vehicle (MRC)
EF2: Trip is interrupted
EF3: Collision risk
1-J. Vehicle can implement FMS
Mandatory Autonomy Initiated Driver in Control Success/Failure
1-D1. There is sufficient time for the driver to react
1-K. Driver detects DDT-fallback is required
1-L. Driver requests take-over
1-M. ADS detects take-over request
1-N. Driver correctly performs driver-initiated take-over DDT
1-O. Driver requests emergency stop
Mandatory Driver Initiated Driver in Control Success/Failure
EF4: Failed hand-over
1-P. Driver correctly performs DDT fallback
Mandatory Autonomy Initiated Driver in Control Success/Failure
1-Q. Driver detects ADS rejection
1-R. Driver correctly performs DDT
Voluntary Driver Initiated Autonomy in Control Failure
yes
no

Figure 13: ESD 1.1 – Driver requests handover.

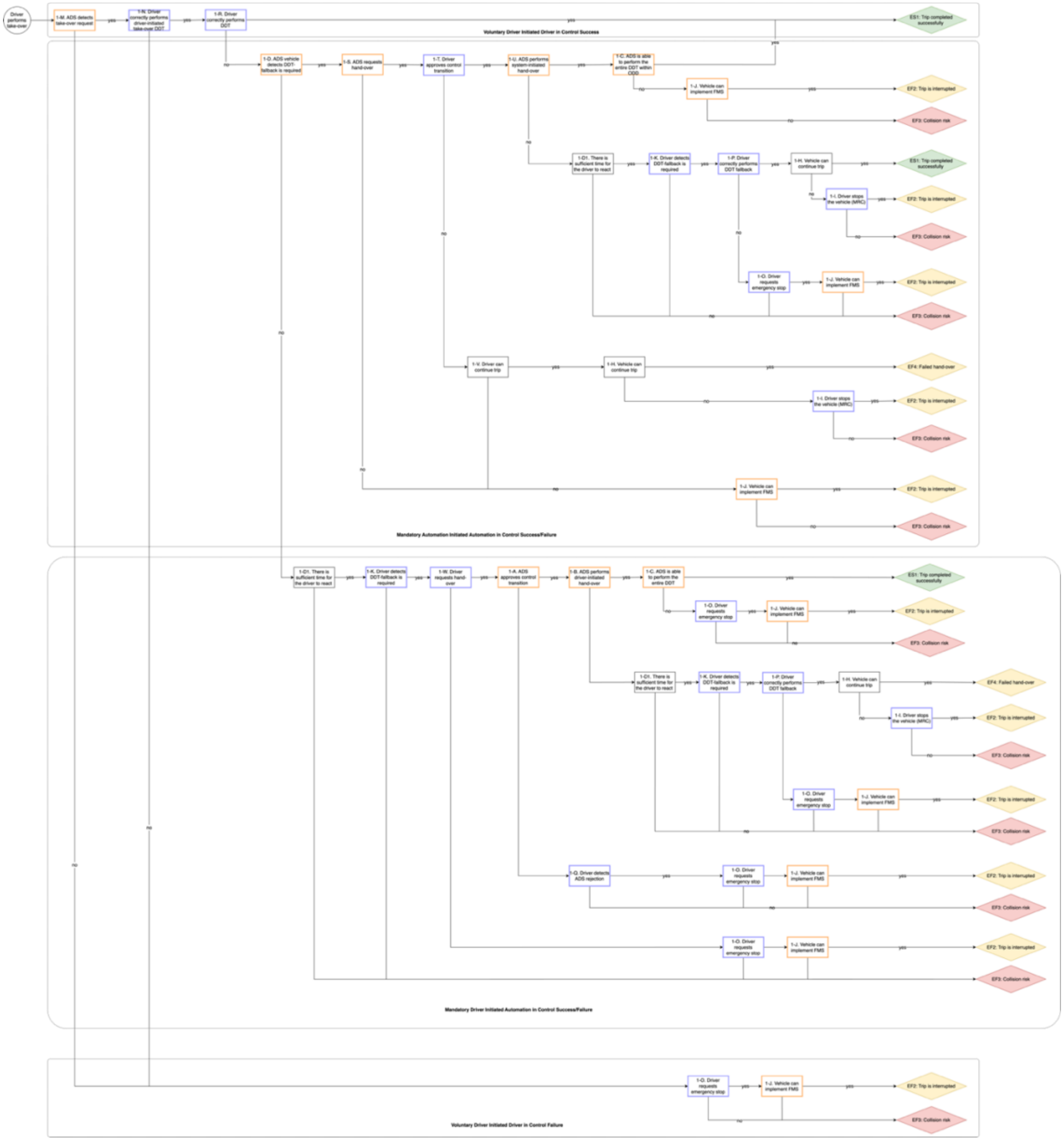

Figure 14: ESD 1.2 – Driver performs takeover.

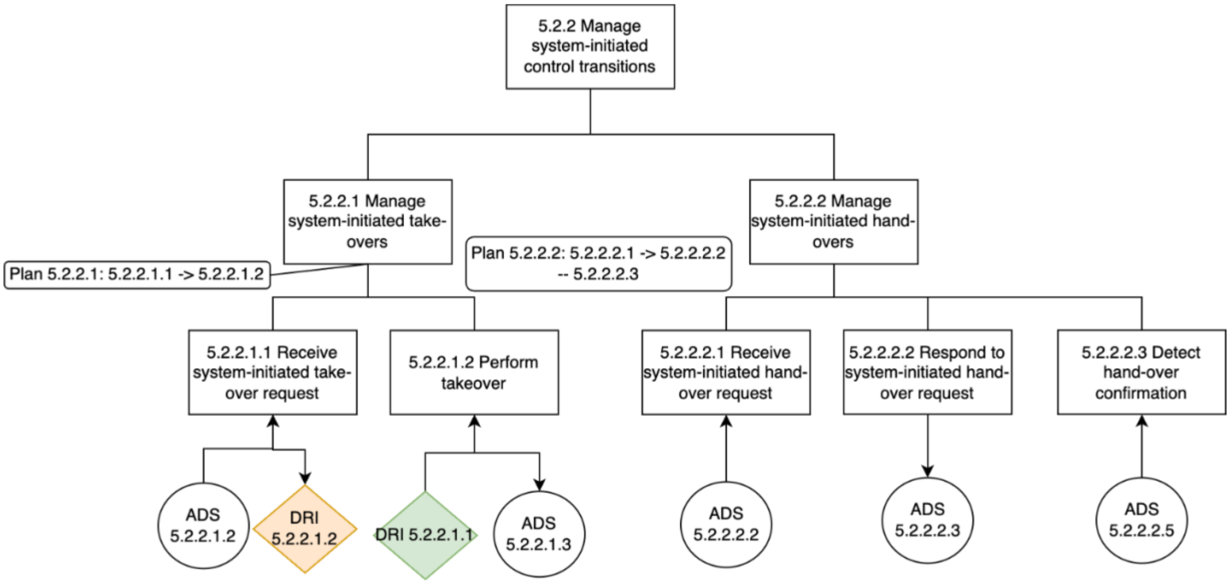

Figure 15: DRI 5.2.2 CoTA.

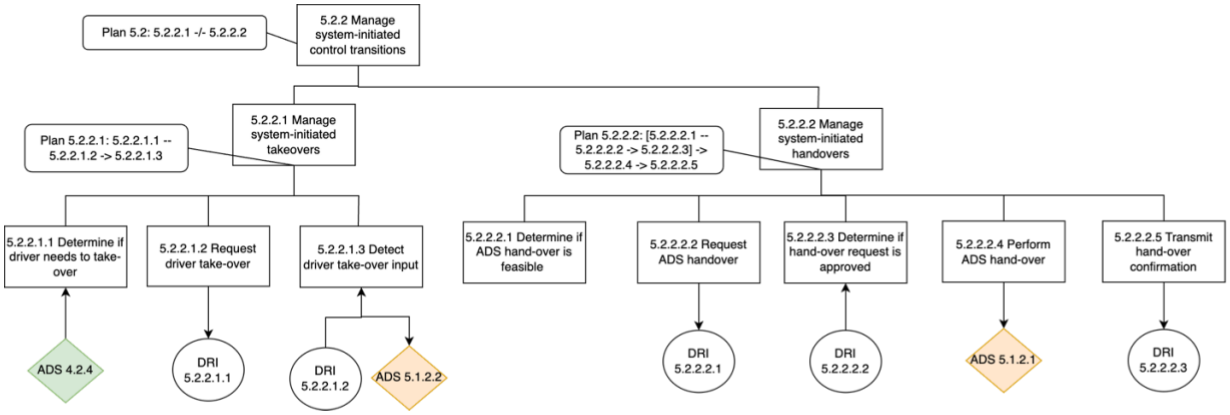

Figure 16: ADS 5.2.2 CoTA.

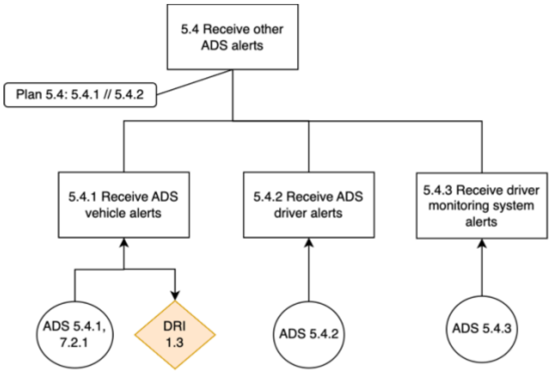

Figure 17: DRI 5.4 CoTA.

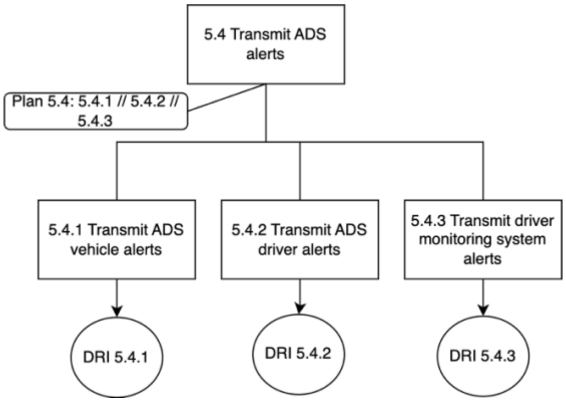

Figure 18: ADS 5.4 CoTA.

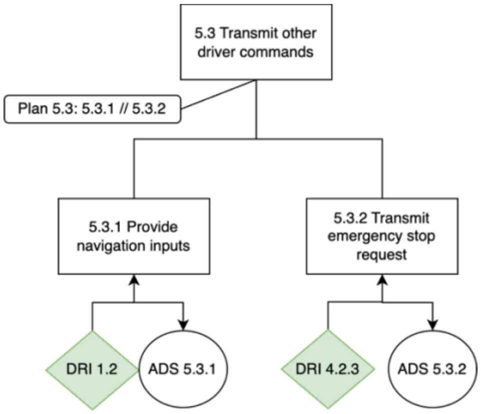

Figure 19: DRI 5.3 CoTA.

Bibliography

Aurora. (2021). *Voluntary Safety Self-Assessment*.

Automated Vehicle Safety Consortium. (2021). *Best Practice for Metrics and Methods for Assessing Safety Performance of Automated Driving Systems (ADS)*. www.sae.org.

Ayas, S., Donmez, B., & Tang, X. (2023). Drowsiness Mitigation Through Driver State Monitoring Systems: A Scoping Review. *Human Factors*. https://doi.org/10.1177/00187208231208523

Bayramova, A., Edwards, D. J., Roberts, C., & Rillie, I. (2023). Constructs of leading indicators: A synthesis of safety literature. *Journal of Safety Research*, *85*, 469–484. https://doi.org/10.1016/j.jsr.2023.04.015

Bhoopalam, A. K., van den Berg, R., Agatz, N., & Chorus, C. G. (2023). The long road to automated trucking: Insights from driver focus groups. *Transportation Research Part C: Emerging Technologies*, *156*. https://doi.org/10.1016/j.trc.2023.104351

Bin-Nun, A. Y., Lizarazo, C., Panasci, A., Madden, S., & Tebbens, R. J. D. (2023). What do surrogate safety metrics measure? Understanding driving safety as a continuum. *Accident Analysis and Prevention*, *195*. https://doi.org/10.1016/j.aap.2023.107245

Bonnefon, J.-François. (2021). *The car that knew too much: Can a machine be moral?* MIT Press.

Camila Correa-Jullian, Marilia Ramos, Ali Mosleh, & Jiaqi Ma. (2024). Task Allocation and Control Transitions in Autonomous Driving System Operations. *34th European Safety and Reliability Conference (ESREL 2024)*.

Cao, Y., Zhou, F., Pulver, E. M., Molnar, L. J., Robert, L. P., Tilbury, D. M., & Yang, X. J. (2021). *Towards standardized metrics for measuring takeover performance in conditionally automated driving: A systematic review*.

Castro, C., Pablo Doncel, P., Ledesma, R. D., Montes, S. A., Daniela Barragan, D., Oviedo-Trespalacios, O., Bianchi, A., Kauer, N., Qu, W., & Padilla, J. L. (2024). Measurement invariance of the driving inattention scale (ARDES) across 7 countries. *Accident Analysis and Prevention*, *195*. https://doi.org/10.1016/j.aap.2023.107412

Chang, Y. H. J., & Mosleh, A. (2007). Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 5: Dynamic probabilistic simulation of the IDAC model. *Reliability Engineering and System Safety*, *92*(8), 1076–1101. https://doi.org/10.1016/j.ress.2006.05.012

Chu, M., Zong, K., Shu, X., Gong, J., Lu, Z., Guo, K., Dai, X., & Zhou, G. (2023, April 19). Work with AI and Work for AI: Autonomous Vehicle Safety Drivers' Lived Experiences. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3544548.3581564

Correa-Jullian, C., Ramos, M., Mosleh, A., & Ma, J. (2024). Operational safety hazard identification methodology for automated driving systems fleets. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*. https://doi.org/10.1177/1748006X241233863

Cosmin-Spanoche, A., Correa-Jullian, C., Xia, X., Mosleh, A., & Ma, J. (2024). Exploring Safety-Related Metrics To Assess Human-System Interactions In Heavy-Duty Automated Vehicles. *34th European Safety and Reliability Conference (ESREL 2024)*.

DeGuzman, C. A., Kanaan, D., Hopkins, S. A., & Donmez, B. (2021). Takeover request (TOR) effects during different automated vehicle failures. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*. https://doi.org/10.1080/15472450.2021.1891536

Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., & Koltun, V. (2017). *CARLA: An Open Urban Driving Simulator*. http://arxiv.org/abs/1711.03938

Ekanem, N. J., Mosleh, A., & Shen, S. H. (2016). Phoenix - A model-based Human Reliability Analysis methodology: Qualitative Analysis Procedure. *Reliability Engineering and System Safety*, *145*, 301–315. https://doi.org/10.1016/j.ress.2015.07.009

Ike. (2019). *Safety Report*.

International Standardization Organization. (2018). *ISO 26262-1:2018 Road vehicles — Functional safety*.

International Standardization Organization. (2022). *ISO 21448:2022 Road vehicles — Safety of the intended functionality*.

Kodiak. (2020). *Kodiak Safety Report 2020*.

Masala, G. L., & Grosso, E. (2014). Real time detection of driver attention: Emerging solutions based on robust iconic classifiers and dictionary of poses. *Transportation Research Part C: Emerging Technologies*, *49*, 32–42. https://doi.org/10.1016/j.trc.2014.10.005

Morales-Alvarez, W., Sipele, O., Léberon, R., Tadjine, H. H., & Olaverri-Monreal, C. (2020). Automated driving: A literature review of the take over request in conditional automation. In *Electronics (Switzerland)* (Vol. 9, Issue 12, pp. 1–34). MDPI AG. https://doi.org/10.3390/electronics9122087

National Highway Traffic Safety Administration. (2017). *Automated Driving Systems 2.0: A Vision for Safety*.

National Highway Traffic Safety Administration. (2023). *Traffic Safety Facts -- 2021 Data: Large Trucks*.

Ramos, M. A., Thieme, C. A., Utne, I. B., & Mosleh, A. (2020a). A generic approach to analysing failures in human – system interaction in autonomy. *Safety Science*, *129*, 104808. https://doi.org/10.1016/J.SSCI.2020.104808

Ramos, M. A., Thieme, C. A., Utne, I. B., & Mosleh, A. (2020b). Human-system concurrent task analysis for maritime autonomous surface ship operation and safety. *Reliability Engineering and System Safety*, *195*. https://doi.org/10.1016/j.ress.2019.106697

Reiman, T., & Pietikäinen, E. (2012). Leading indicators of system safety - Monitoring and driving the organizational safety potential. *Safety Science*, *50*(10), 1993–2000. https://doi.org/10.1016/j.ssci.2011.07.015

SAE International. (2021). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. *SAE Standard J3016 APR2021*.

Sheehan, C., Donohue, R., Shea, T., Cooper, B., & De Cieri, H. (2016). Leading and lagging indicators of occupational health and safety: The moderating role of safety leadership. *Accident Analysis and Prevention*, *92*, 130–138. https://doi.org/10.1016/j.aap.2016.03.018

Shetty, A., Tavafoghi, H., Kurzhanskiy, A., Poolla, K., & Varaiya, P. (2022). Automated Vehicle Safety and Deployment: Lessons from Human Crashes. *2022 International Conference on Connected Vehicle and Expo (ICCVE)*, 1–6. https://doi.org/10.1109/ICCVE52871.2022.9742994

Talebian, A., & Mishra, S. (2022). Unfolding the state of the adoption of connected autonomous trucks by the commercial fleet owner industry. *Transportation Research Part E: Logistics and Transportation Review, 158,* 102616. https://doi.org/10.1016/J.TRE.2022.102616

Torc. (2021). *Innovating Safety and Efficiency - Torc Safety Report*.

TuSimple. (2020). *TuSimple Safety Report Version 2.0*.

UL Standards and Engagement. (2023). *Standard for Safety for the Evaluation of Autonomous Products, UL 4600*.

Wang, C., Xie, Y., Huang, H., & Liu, P. (2021). A review of surrogate safety measures and their applications in connected and automated vehicles safety modeling. *Accident Analysis and Prevention*, *157*. https://doi.org/10.1016/j.aap.2021.106157

Wishart, J., Como, S., Elli, M., Russo, B., Weast, J., Altekar, N., & James, E. (2020). Driving Safety Performance Assessment Metrics for ADS-Equipped Vehicles. *SAE Technical Papers*, *2020-April*(April). https://doi.org/10.4271/2020-01-1206

Xu, R., Guo, Y., Han, X., Xia, X., Xiang, H., & Ma, J. (2021). OpenCDA: An Open Cooperative Driving Automation Framework Integrated with Co-Simulation. *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 1155–1162. https://doi.org/10.1109/ITSC48978.2021.9564825

Xu, R., Xiang, H., Han, X., Xia, X., Meng, Z., Chen, C. J., Correa-Jullian, C., & Ma, J. (2023). The OpenCDA Open-Source Ecosystem for Cooperative Driving Automation Research. *IEEE Transactions on Intelligent Vehicles*, *8*(4), 2698–2711. https://doi.org/10.1109/TIV.2023.3244948

Yang, X. J., Pradhan, A. K., Tilbury, D., & Robert, L. (2018). *Human Autonomous Vehicles Interactions: An Interdisciplinary Approach*.

Zhai, S., Gao, S., Wang, L., & Liu, P. (2023). When both human and machine drivers make mistakes: Whom to blame? *Transportation Research Part A: Policy and Practice*, *170*, 103637.

Zhang, B., de Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *64*, 285–307. https://doi.org/10.1016/j.trf.2019.04.020