SAFETY ASSURANCE FRAMEWORK FOR CONNECTED, AUTOMATED MOBILITY SYSTEMS

D2.3

Final SUNRISE Safety Assurance Framework

Project short name
SUNRISE

Project full name
Safety assUraNce fRamework for connected, automated mobIlity SystEms

Horizon Research and Innovation Actions | Project No. 101069573
Call HORIZON-CL5-2021-D6-01

Funded by
the European Union

ccam-sunrise-project.eu/

Dissemination level	Public (PU) - fully open
Work package	WP2: CCAM safety assurance framework
Deliverable number	D2.3: Final SUNRISE safety assurance framework
Deliverable responsible	Jason Xizhe Zhang, University of Warwick
Status - Version	Final – V1.0
Submission date	27/08/2025
Keywords	SUNRISE, SAF, Safety Assurance Framework, CCAM

Authors/Contributors

Role	Name
Main authors:	Jason Xizhe Zhang (University of Warwick)
Contributing authors:	Olaf Op den Camp (TNO), Stefan de Vries (IDIADA), Bryan Bourauel (BASt), Bernhard Hillbrand (Virtual Vehicle Research), Marcos Nieto Doncel (Vicomtech), John-Fredrik Gronvall (Chalmers), Darko Stern (AVL), Anastasia Bolovinou (ICCS), Ainhoa Arrieta Fernández (ERTICO), Jobst Beckmann (ika)

Quality Control

	Name	Organisation	Date
Peer review 1	Stefan de Vries	IDIADA	11/07/2025
Peer review 2	Bryan Bourauel	BASt	23/07/2025

Version history

Version	Date	Author	Summary of changes
0.1	20/12/2023	Jason Zhang	First draft of document structure.
0.2	27/12/2023	Olaf Op den Camp	Overview of NATM pillars Scenario vs. test scenario. Textual definition of the SAF
0.8	01/04/2025	Jason Xizhe Zhang	Merged sections input, created the first mature version
0.9	22/07/2025	Jason Xizhe Zhang	Addressed internal review comments
1.0	28/08/2025	Jason Xizhe Zhang	Final version

Legal disclaimer

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

Abbreviation	Meaning
AB	Advisory Board
AD	Automated Driving
ADS	Automated Driving System
AEB	Autonomous Emergency Braking
ALKS	Automated Lane Keeping System(s)
AQCG	Automated Query Criteria Generation
ASAM	Association for Standardization of Automation and Measuring Systems
ASE	(Not explicitly defined in the document, appears to be a database format)
AVSC	Automated Vehicle Safety Consortium
CCAM	Connected, Cooperative and Automated Mobility
COTSATO	COncretizing Test Scenarios and Associating Test Objectives
CRHS	Cooperative Resilience and Healing System
DDT	Dynamic Driving Task
DF	Data Framework
EC	European Commission
EP	Expert Platform
FMI	Functional Mock-up Interface
FuSa	Functional Safety
GAMAB	Globalement Au Moins Aussi Bon (Generally At Least As Good)
GRVA	Working Party on Automated/Autonomous and Connected Vehicles
GSN	Goal Structuring Notation
GSR	General Safety Regulation
HMI	Human Machine Interface
IB	Industry Board
ISMR	In-Service Monitoring and Reporting
ISO	International Organization for Standardization
IdP	Identity Provider

JRC	Joint Research Center
KATRI	Korea Automotive Technology Institute
KPI	Key Performance Indicator(s)
NATM	New Assessment/Test Method for Automated Driving
ODD	Operational Design Domain
OEM	Original Equipment Manufacturer
OSI	Open Simulation Interface
PDF	Probability Density Function
PRB	Positive Risk Balance
SACP	Situational Awareness and Collaborative Perception
SAE	Society of Automotive Engineers
SAF	Safety Assurance Framework
SCDB	SCenario DataBase
SMS	Safety Management System
SOTA	Secure Over-the-Air
SOTIF	Safety Of The Intended Functionality
SRI	Scenario Relatedness Index
STPA	Systems-Theoretic Process Analysis
SUT	System Under Test
TDMS	Trust Data Management System
TRL	Technology Readiness Level
UC	Use Case
UI	User Interface
UNECE	United Nations Economic Commission for Europe
V&V	Verification and Validation
V2X	Vehicle-to-Everything
VSOC	Vehicle Security Operations Center
WP	Work Package
XiL	X-in-the-Loop

EXECUTIVE SUMMARY

Safety assurance of Cooperative, Connected, and Automated Mobility (CCAM) systems is a crucial factor for their successful adoption in society, yet it remains a significant challenge. It is generally acknowledged that for higher levels of automation, the validation of these systems by conventional test methods would be infeasible. Furthermore, certification initiatives worldwide struggle to define a harmonized safety assurance approach enabling massive deployment of CCAM systems.

The **SUNRISE** project develops and demonstrates a **CCAM Safety Assurance Framework** (SAF). The overall objective of the SUNRISE project is to accelerate the large-scale and safe deployment of CCAM systems. In alignment with international twin projects and initiatives, the project aims to achieve this objective by providing a SAF consisting of three main components: a Method, a Toolchain and a Data Framework. The **Method** is established to support the SAF safety argumentation, and includes procedures for scenario selection, sub-space creation, (dynamic) allocation of tests to test instances and the definition of a variety of metrics and rating procedures. The **Toolchain** contains a set of tools for safety assessment of CCAM systems, including approaches for virtual, hybrid and physical testing. The **Data Framework** provides online access, connection and harmonization of external Scenario Databases (SCDBs), allowing its users to perform query-based extraction of safety relevant scenarios, allocation of selected scenarios to a variety of test environments, and reception of the test results.

The SUNRISE SAF aligns with the New Assessment/Testing Method (**NATM**) as proposed by WP 29 of United Nations Economic Commission for Europe (**UNECE**) [1]. With the SAF, SUNRISE provides the operationalization of the NATM approach, making it valuable for all stakeholders, including certifiers, regulators and industry. The SAF is designed to support a structured, interoperable, scalable safety assurance framework for CCAM systems. This document starts with the SAF fundamentals, together with its link to policies, standards, regulation and legislations, and the introduction of the concept of scenarios. Followed by the explanations on how the SAF enables the querying of scenario databases from the input layer, the concretization and generation of test scenarios, the allocation of tests to test instances, the test execution, and the analysis of the results in a way that aligns with regulatory needs and practical industry requirements. It then dives into the SUNRISE Data Framework design and implementation, and its position within the SUNRISE SAF. Detailed summaries of many specific topics discussed during the project duration are presented, such as the SAE level covered, the link between SUNRISE SAF and the role of the human driver, cybersecurity, and being future proof. Lastly, SAF application guidelines are provided to cater its target users, together with a listing of its current limitations.

1 GENERAL INTRODUCTION

1.1 Project introduction

Safety assurance of CCAM systems is a crucial factor for their successful adoption in society, yet it remains a significant challenge. CCAM systems need to demonstrate reliability in all driving scenarios, requiring robust safety argumentation. It is acknowledged that for higher levels of automation, the validation of these systems by means of real test-drives would be infeasible. In consequence, a carefully designed mixture of physical and virtual testing has emerged as a promising approach, with the virtual part bearing more significant weight for cost efficiency reasons.

Worldwide, several initiatives have started to develop test and assessment methods for Automated Driving (AD) functions. These initiatives already transitioned from conventional validation to a scenario-based approach and combine different test instances (physical and virtual testing) to avoid the million-mile issue.

The initiatives mentioned above, provide new approaches to CCAM validation, and many expert groups formed by different stakeholders, are already working on CCAM systems' testing and quality assurance. Nevertheless, the lack of a common European validation framework and homogeneity regarding validation procedures to ensure safety of these complex systems, hampers the safe and large-scale deployment of CCAM solutions. In this landscape, the role of standards is paramount in establishing common ground and providing technical guidance. However, standardising the entire pipeline of CCAM validation and assurance is in its infancy, as many of the standards are under development or have been very recently published and still need time to be synchronised and established as common practice.

Scenario Databases (SCDBs) are another issue tackled by several initiatives and projects, that generally tends to silo solutions. A clear concrete approach should be used (at least at European level), dealing with scenarios of any possible variations, including the creation, editing, parameterisation, storing, exporting, importing, etc. in a universally agreed manner.

Furthermore, validation methods and testing procedures still lack appropriate safety assessment criteria to build a robust safety case. These must be set and be valid for the whole parameter space of scenarios. Another level of complexity is added, due to regional differences in traffic rules, signs, actors and situations.

Evolving from the achievements obtained in HEADSTART and taking other project initiatives as a baseline, it becomes necessary to move to the next level in the development and demonstration of a commonly accepted **SAF** for the safety validation of CCAM systems, including a broad portfolio of Use Cases (UCs) and comprehensive test and validation tools. This will be done in **SUNRISE**, which stands for **S**afety ass**U**ra**N**ce f**R**amework for connected, automated mob**I**lity **S**yst**E**ms.

The SAF is the main product of the SUNRISE project. As Figure 1 indicates, it takes a central role, fulfilling the needs of different automotive stakeholders that all have their own interests in using it.

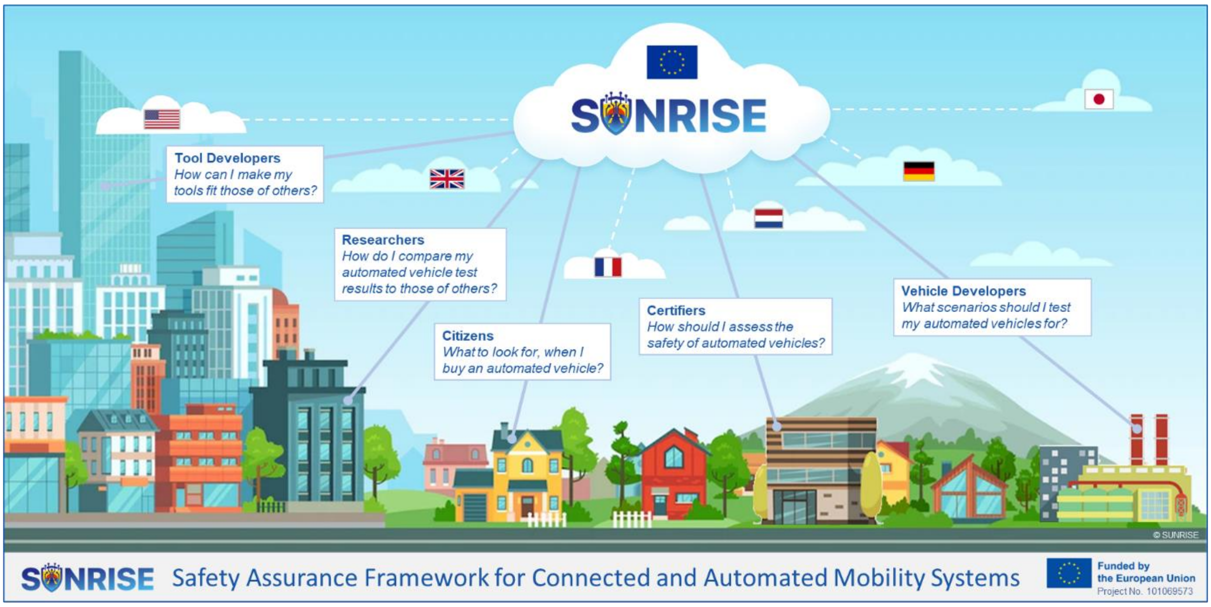

Figure 1: Safety Assurance Framework stakeholders

The **overall objective** of the SUNRISE project is to accelerate the safe deployment of innovative CCAM technologies and systems for passengers and goods by creating demonstrable and positive impact towards safety, specifically the EU's long-term goal of moving close to zero fatalities and serious injuries by 2050 (Vision Zero), and the resilience of (road) transport systems. The project aims to achieve this objective by providing a SAF consisting of three main components: a Method, a Toolchain and a Data Framework. The **Method** is established to support the SAF safety argumentation, and includes procedures for scenario selection, sub-space creation, dynamic allocation to test instances and a variety of metrics and rating procedures. The **Toolchain** contains a set of tools for safety assessment of CCAM systems, including approaches for virtual, hybrid and physical testing. The **Data Framework** provides online access, connection and harmonization of external SCDBs, allowing its users to perform query-based extraction of safety relevant scenarios, allocation of selected scenarios to a variety of test environments, and generation of the test results. The SAF will be put to the test by a series of **Use Cases demonstrations**, designed to identify and solve possible errors, gaps and improvements to the underlying methods, tools and data.

Following a common approach will be crucial for present and future activities regarding the testing and validation of CCAM systems, allowing to obtain results in a standardised way, to improve analysis and comparability, hence maximising the societal impact of the introduction of CCAM systems.

Figure 2 shows the general workplan of the SUNRISE project.

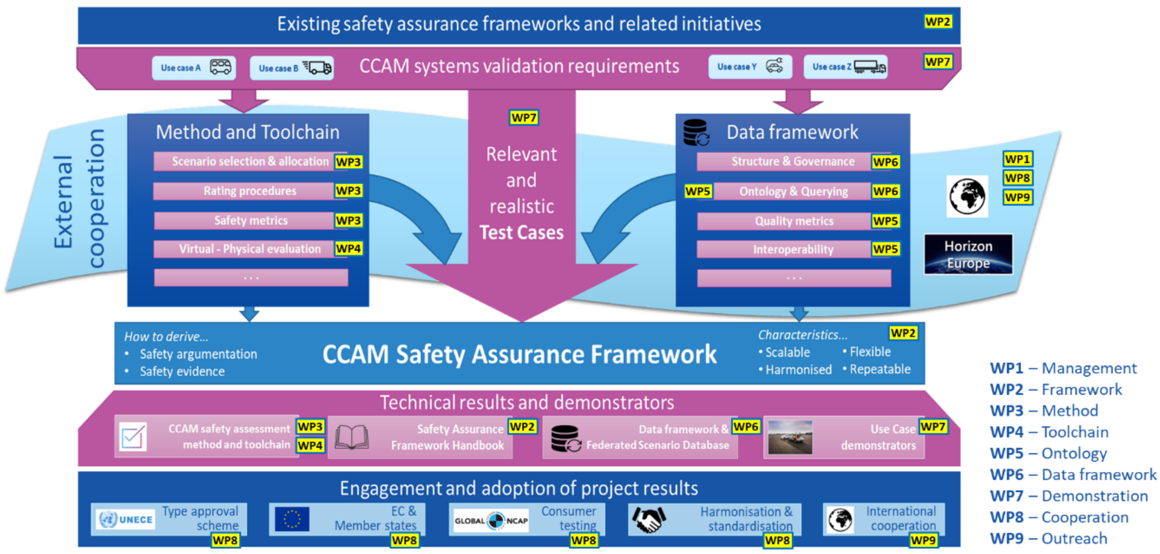

Figure 2: Workplan of the SUNRISE Project

1.2 Purpose of deliverable

The purpose of this deliverable is to present the final version of the SAF developed by the SUNRISE project. The SAF is designed to support a structured, interoperable, scalable safety assurance workflow for CCAM systems. This document explains how the SAF enables the querying of scenario databases, the concretization and generation of test scenarios, the allocation of tests to test instances, the test execution, and the analysis of the results in a way that aligns with regulatory needs and practical industry requirements.

This deliverable consolidates the SAF's structure and components, drawing on existing standards and validation frameworks from both international and European initiatives. It also incorporates feedback from certification authorities, vehicle manufacturers, standardisation organisations, and other EU projects. As such, the SAF described in this document is the result of iterative refinement based on both technical development and external stakeholder engagement. The document sets out the SAF’s functional elements, including how scenarios are accessed and managed, how tests are carried out in different environments, and how evidence is gathered and analysed to support safety assessment. It also explains how the SAF accommodates ongoing monitoring during vehicle deployment and how it can be applied by different types of users, including certifiers and industry.

This deliverable serves as a harmonised foundation for understanding and applying the SAF, including insight on how the framework was designed and how it aligns with external regulations and projects. Although the SAF is part of a broader set of SUNRISE results, this document is intended to be self-contained and comprehensive, describing the SAF in full without requiring reference to other deliverables. It therefore functions as the principal definition and documentation of the SAF at the conclusion of the SUNRISE project.

1.3 Intended audience

The SUNRISE SAF is intended for a diverse set of stakeholders engaged in the development, validation, regulation, and certification of CCAM systems. These users can be broadly categorised into two groups: **commercial users**, such as OEMs and Tier 1 suppliers, and **non-commercial users**, including international vehicle safety bodies and regulators.

Commercial users are primarily concerned with ensuring the safe integration and validation of CCAM systems while managing cost, time, and regulatory complexity. Their key challenges include the absence of a unified safety validation approach, limited access to high-quality scenario databases, inconsistent validation tools and methods across organisations, and the high cost and time burden associated with meeting regulations and standards such as UNECE WP.29 [1], ISO 26262 [2], and SOTIF [3]. For these users, the SAF provides a standardised and scalable methodology that improves efficiency and reliability in testing, enhances coverage through harmonised access to scenario databases, and enables cost-effective simulation-based verification and validation workflows. It also supports accelerated compliance by enabling iterative refinement of safety requirements through structured feedback loops.

Non-commercial users, such as certification entities, type approval authorities and policymakers, face their own set of issues, including the lack of standardised testing methodologies, gaps in scenario coverage, slow adaptation to emerging technologies, and high costs associated with complex compliance evaluations. The SAF addresses these pain points by offering a harmonised assessment methodology aligned with international efforts (e.g. NATM [1]), enabling data-driven policy making and scalable compliance testing. It also supports future-proof regulation updates through the use of real-world data from SCDBs and in-service monitoring. Additionally, the SAF promotes transparency, public trust, and international alignment in safety evaluations.

1.4 Deliverable structure and relation to other parts of project

This deliverable presents the final version of the SUNRISE SAF, structured to reflect its alignment with project-wide activities, stakeholder needs, and relevant standards and regulations. It builds upon foundational work conducted in Task 2.1 (as reported in deliverable D2.1 [4]), where various safety assurance frameworks were benchmarked and compared. The key findings from that benchmarking exercise was input into the development process of the SAF.

Chapter 2 introduces the fundamentals of the SAF, including its alignment with project goals, regulatory bodies, consumer testing organisations, standardisation communities, and other EU-funded projects such as SELFY and i4Driving. Chapters 3 and 4 detail the core architecture of the SAF, particularly the Performance Assurance Workflow (Scenario, Environment, Safety Argument), presenting its structure, roles, and interrelations. Chapter 5 outlines the SUNRISE Data Framework, which enables interoperable and traceable management of scenarios. Chapter 6 explores how the SAF addresses specific priority topics such as SAE levels, human involvement, cybersecurity, unknown scenarios, and Artificial Intelligence. Chapter 7 provides detailed guidance for applying the SAF in practice, covering all SAF blocks. Chapter 8 discusses current limitations of the SAF, while Chapter 9 concludes the deliverable with reflections on its significance and outlook for future work.

The SAF explained in this deliverable serves as the central element connecting all SUNRISE work packages. Work Packages WP3-6 cover the development of various components within the SAF. For example, WP3 focuses on concretisation and test instance allocation, WP4 focuses on virtual testing framework, WP5 focuses on Data Framework requirements, operations and assessment metrics, WP6 focuses on the development of the SUNRISE Data Framework. In addition, WP7 demonstrates and validates the complete SAF, together with the developed components, by means of various use cases. And WP8 and WP9 focus on the stakeholder engagement and feedback.

Feedback from WP7 (about SAF demonstration and validation) was exchanged purposefully at the work package and task level, allowing the input from the use case demonstrations to be directly mapped to the responsible development teams. In particular, insights gained from WP7 task T7.3, which were derived from the eight practical use cases, were promptly integrated into the relevant development work packages and taken into account during the further advancement of the SAF. Continuous alignment was ensured through regular discussions regarding deliverables and ongoing communication among all participating partners. Additionally, the leads of the relevant SAF work packages and tasks were actively involved in WP7 and Task 7.3. This approach ensured effective knowledge transfer and made certain that practical experiences and challenges resulting from WP7, were systematically incorporated into the validation and optimization of the SAF in WP3-6.

2 SAF INTRODUCTION

SUNRISE stands for Safety assUraNce fRamework for connected and automated mobIlity SystEms. The main **goal** of the SUNRISE project is to develop a harmonized **CCAM** SAF that fulfils the needs of different **users** including regulators, certifiers, manufacturers (OEMs) and their suppliers (TIER1s). The following concept image provides a high-level overview of the SUNRISE Safety Assurance Framework and its key components.

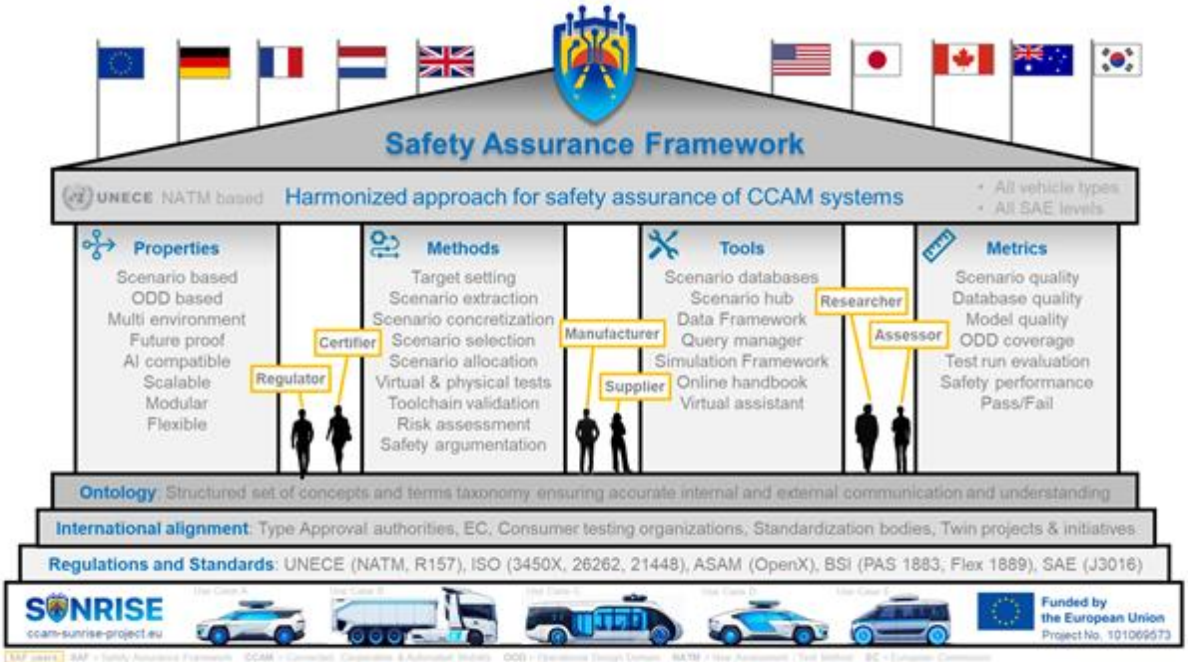

Figure 3: Concept image of the SUNRISE Safety Assurance Framework

The SAF can be applied to **all vehicle types** (for public road use) and **all SAE levels**. The SAF **features** the use of scenarios and the Operational Design Domain (ODD) [5] as the basis for safety assurance of CCAM systems. Test scenarios can be executed in multiple test environments, ranging from purely virtual to purely physical. Moreover, the SAF is future proof. This means that future technologies and scenarios can also be handled (like AI based systems), and that it easily adapts to new standards and regulations.

The SAF is built on **3 main pillars**: Methods, Tools and Metrics. The **Methods** pillar includes for example a method for target setting, scenario selection and toolchain validation. The **Tools** pillar covers a scenarios hub (also called the Data Framework), a query manager to find scenarios, a V&V simulation framework and an online SAF Handbook with a virtual assistant. Finally, the **Metrics** pillar includes for example metrics on scenario and database quality, ODD coverage, safety performance and pass/fail criteria.

The SAF is based on various **standards and regulations**, most importantly the UNECE NATM [1]. Furthermore, the SAF has been **aligned with international projects, initiatives and entities** operating in the field of CCAM safety assurance. For example with type approval entities (such as UNECE, KBA [DE] and RDW [NL]), consumer testing agencies (such as Euro

NCAP [6]), standardization bodies (such as ISO [7], ASAM [8] and BSI [9]) and with twin projects and initiatives (such as VVM [DE] [10], Sakura [JP] [11] and Assess [KR]). Besides that, the Data Framework connects to a variety of international scenario databases (such as StreetWise [NL] [12], Safety Pool™ scenario database [UK] [13], ADScene [FR] [14], scenario.center [DE] [15] and AVL Scenius [AT] [16]). The foundations of the SAF also count with an **ontology** to ensure accurate information exchange among its components, and the SAF has been validated and demonstrated by a large variety of **Use Cases**, represented by the bottom layer in Figure 3 above.

As mentioned above, the SUNRISE SAF is based on the UNECE NATM. It includes an audit of the safety assessment process, the safety management system as followed by the CCAM manufacturer, and the used test methods and test tools. The framework adopts a scenario-based approach, where tests (virtual testing, track testing and real-world testing) are based on scenarios taking into account the CCAM system's ODD for assessing the safety of the system. Additionally, it provides procedures to monitor that the CCAM's safety is maintained during the system's lifecycle. The scenarios that feed into the SAF are knowledge based and/or data driven.

The SAF includes, but is not limited to, processes to:

- generate the relevant scenarios,
- allocate test scenarios to the different test methods,
- execute the test scenarios,
- assess and analyse the test results to come to a statement about safety assurance, and
- monitor the CCAM system and provide information of the system to the stakeholders to ensure safety during deployment.

Some of the key characteristics of the SAF include:

- It is a framework that can be seen as a structure of processes and procedures. A *process* is a series of tasks and activities that produce an outcome whereas a *procedure* is a set of instructions for completing a single task or activity.
- The framework is harmonized, such that the use of the framework and hence the outcome is accepted by all stakeholders.
- The SAF is dedicated to get an understanding whether a CCAM system meets a certain predefined level of safety.
- The safety assessment does not stop at the introduction of the CCAM system into operation (usually the road), but continues to provide information on the safety performance of the system during deployment.

Thereafter a more **detailed insight** is provided in the following sections, with a reference to the UNECE NATM multi-pillar approach. The different pillars in the SAF will be mentioned: the audit pillar, the performance assurance pillars, and the monitoring and reporting pillar. Moreover, it is indicated that the SAF is scenario-based, meaning that tests are based on scenarios (data-driven and/or knowledge based [17]) that sufficiently cover the ODD of the CCAM system. Note: the fact that processes are considered part of the SUNRISE SAF does not mean automatically that the SUNRISE project also considers all of these processes within the scope of the project. This is decided outside SUNRISE task T2.2.

The pillars supporting the safety assessment of CCAM systems is illustrated in Figure 4. These pillars form the basis of the SUNRISE SAF as shown below.

Figure 4: SUNRISE SAF based on Pillars of the UNECE NATM document

As shown in Figure 4 above, the SAF consists of four key colour-coded components: blue, blue-to-red, yellow, and purple. These represent the Audit, Safety Performance Assurance, In-Service Monitoring and Reporting (ISMR), and Input elements of the framework, respectively. This structure aligns with the multi-pillar approach developed by the UNECE WP.29 Working Party on Automated/Autonomous and Connected Vehicles (GRVA), as outlined in the New Assessment/Test Methods (NATM) Master Document [1]. The NATM, adapted and elaborated by Donà et al. [18] in collaboration with Uni Systems Italy and the European Commission Joint Research Centre (JRC), provides a scenario-based certification framework composed of several pillars. These include three distinct testing methods—real-world field operational tests, proving ground evaluations (such as Euro NCAP), and virtual simulation. Virtual simulation, in particular, has emerged as a critical component for enabling large-scale, scenario-based safety assessments.

Within this broader structure, the SAF Input layer (purple) addresses foundational requirements related to the ODD, system behaviour, and other external factors and test objectives that inform safety validation. The yellow section, representing ISMR, captures operational data during deployment and large-scale testing. This pillar supports continuous safety assessment and improvement by monitoring for erratic system behaviour, performance degradation, or unanticipated scenarios. It also enables scenario sharing and feedback mechanisms among regulators and manufacturers. The ISMR phase therefore plays a dual role in both monitoring and reporting.

The Audit pillar (blue) focuses on evaluating the manufacturer's Safety Management System (SMS), the safety assessment process, and the reliability of tools and environments used for testing (incl. used scenario databases). This analytical phase ensures that safety assurance is not limited to test results alone but includes the robustness of the underlying development and assessment processes. It confirms that manufacturers have appropriate procedures in place throughout the system's lifecycle, covering both operational and functional safety.

Although the primary focus of the SUNRISE project lies within the Safety Performance Assurance pillar (blue-to-red), the Audit and ISMR pillars remain integral to the overarching structure of the SAF. They provide essential context and continuity for assessing system safety over time, in line with the vision set forth in the NATM framework for automated vehicle certification.

2.1 SAF fundamentals and alignment

This section outlines the foundations of the SUNRISE SAF and demonstrates how it has been developed in close alignment with key international initiatives, regulatory bodies, standardisation organisations, EU projects, and individual CCAM experts. Building on extensive analysis of existing safety assurance approaches and active engagement with global twin projects and initiatives, the SAF offers a harmonised and coherent structure for CCAM safety validation. The purpose of this section is to present the international harmonization behind the SAF, describe its development process, and highlight the extensive cross-project, cross-border, and cross-sector alignment that underpins its credibility and applicability.

2.1.1 SAF Fundaments

2.1.1.1 Findings from T2.1

Evaluation effort was firstly carried out in Task 2.1. The results have been documented in SUNRISE deliverable, D2.1 *"Overview and Gaps of Existing Safety Assessment Frameworks"*. It provided a detailed comparative analysis of safety assurance methods applicable to automated driving systems. It reviewed a selection of international and national initiatives, including the UNECE NATM [1], PEGASUS [19], HEADSTART [20], StreetWise [12], SAKURA [11], Safety Pool™ scenario database [13], ADScene [14], and ArchitectECA2030 [21].

The analysis identified several important findings and limitations across existing frameworks:

- **Scenario Coverage**: Most frameworks lack systematic methods for representing and reusing scenario data across simulation, track testing, and real-world validation. While PEGASUS and SAKURA propose structured scenario pipelines, their scenario definitions are often specific to national contexts and not readily generalisable or interoperable.

- **Tool Interoperability**: There is limited integration across tools used for scenario creation, execution, and safety argumentation. This leads to fragmented assurance workflows and restricts the traceability of safety claims.

- **Handling of Uncertainty and Edge Cases**: Although some frameworks address uncertainty qualitatively, few provide quantitative methods for bounding risk from unknown scenarios or low-likelihood events. Formal mechanisms for incorporating domain gaps or distribution shifts are often missing.

- **Performance Criteria and Decision Rules**: Clear pass/fail criteria are often undefined or left to manufacturer interpretation. There is limited guidance on how performance thresholds should be justified or linked to ODDs.

- **In-Service Monitoring Integration**: So far, most frameworks concentrate on pre-deployment testing. The SUNRISE project has not focused on a structured linkage between in-use data collection and the original safety assurance argument. In the CERTAIN project, however, such a link is established, making use of the scenario-based approach that has been developed in SUNRISE.

- **Federated Data Handling**: Few frameworks address how scenario data and safety evidence can be managed across distributed, multi-stakeholder environments. This creates challenges for scalability and transparency in regulatory processes.

These findings directly derived into a set of requirements for the development of the SAF. By integrating these elements, the SAF aims to bridge the gaps outlined in D2.1 and offer a more coherent and practical basis for safety assurance across both development and deployment stages.

2.1.1.2 Survey carried out in T2.2

As a starting point to creating the SUNRISE SAF, within Task 2.2 effort was undertaken to list and investigate existing CCAM safety assurance initiatives across the ecosystem with publicly available information. They include: EU ADS Act [22], SAKURA [11], StreetWise [12], FRAV VMAD [23], VVM [10], and CertiCAV [24]. This section presents a consolidated overview of existing safety assurance initiatives for CCAM systems, focusing on their scenario methodologies, validation strategies, and coverage metrics.

1. Approaches to Safety Assurance

During the survey process, for each of the subjects the following questions are asked:

1) How are scenarios derived? Covering creation method, scenario format, scenario storage

2) How is simulation used? How is it validated?

3) What/where/how much real-world testing?

4) What metrics were used?

5) What format was the Safety Argument created in?

The reviewed frameworks generally align with a scenario-based validation philosophy but diverge in emphasis, granularity, and methodological rigour. The **EU ADS Regulation** centres on regulatory compliance, offering a formalized set of predefined use cases (e.g., hub-to-hub transport, valet parking, yard automation) and mandates that manufacturers assess functional and operational safety across nominal, critical, and failure conditions. It proposes the use of both knowledge-based and data-driven scenario generation strategies, tailored to the declared ODD of the CCAM system.

The **SAKURA project** (Japan) also adopts a scenario-based methodology but grounds its framework in a "Physical Principles Approach." This perspective decomposes driving tasks into Perception, Judgement, and Operation domains, linking scenario construction directly to sensor limits and physics-based interaction models. Scenario generation is structurally categorized into perception disturbances, traffic disturbances, and vehicle motion disturbances. This results in a finite yet holistic set of scenarios, reflecting real-world risks grounded in physical system behaviour.

StreetWise, developed by TNO, builds its framework around large-scale real-world data collection. Unlike generative methods, it identifies and parameterizes scenarios from sensor-level driving data. The database comprises over 45,000 concrete scenarios categorized into 10 common highway cases, emphasizing representativity and statistical coverage. This empirical grounding supports probabilistic risk estimation, especially useful for calculating residual safety risk based on exposure, probability of failure, and severity.

FRAV/VMAD, under the UN's GRVA NATM initiative, emphasizes a multi-pillar approach integrating development audits, proving ground tests, and in-service monitoring. It recommends virtual testing for unsafe-to-reproduce edge cases and promotes scenario generation via three routes: data-driven, knowledge-driven, and concept-based (e.g., recombination of existing scenarios). Despite the maturity of its conceptual framework, specifics on scenario formatting and validation metrics remain under discussion.

VVMethods, coordinated by ika, introduces a layered abstraction model using the OMEGA and ASE formats for structured scenario extraction. It leverages a hybrid scenario generation method combining traffic data with logic-based decomposition. Logical scenarios are modularized into Enveloping, Base, and Focus scenarios, which are stored in a PostgreSQL-based ASE database and translated into OpenX simulation instructions. Criticality assessment, parameter sampling, and risk modelling are embedded throughout the framework.

CertiCAV, emerging from UK research, highlights the centrality of the ODD in shaping scenario relevance, behaviour libraries, and safety arguments. It promotes scenario completeness through a systematic mapping of qualitative, logical, and concrete scenarios, with a strong link to expected behaviour under specified assumptions. The framework reflects a broader policy orientation by offering a structure for simulation evidence generation within national certification contexts.

2. Scenario Generation and Storage

Across initiatives, scenario generation techniques fall into three main categories: real-world identification (StreetWise), systematic derivation from functional models (SAKURA, VVMethods), and regulatory-driven enumeration (EU Regulation, VMAD). StreetWise uniquely emphasizes empirical frequency (exposure) and relies on high-fidelity logging, while VVMethods uses expert-informed sampling to fill in underrepresented edge conditions.

Scenario storage solutions vary from loosely defined repositories (EU Regulation, VMAD) to tightly integrated SQL databases (VVMethods). Open formats such as OpenSCENARIO [25] and OpenDRIVE [26] are common output formats across frameworks, facilitating reuse and interoperability in simulation environments.

3. Simulation and Validation Strategies

Simulation plays a critical role in all reviewed frameworks. SAKURA and VVMethods provide explicit tool validation methods, comparing simulated and real-world steady-state behaviour. VMAD encourages simulation especially for high-risk or low-probability scenarios, though toolchain validation is left to manufacturers. CertiCAV and StreetWise stress that simulation should be grounded in the declared ODD, and validated through consistent coverage and behavioural fidelity.

Most frameworks agree that simulation cannot stand alone—real-world data, whether for exposure estimation (StreetWise), validation sampling (VMAD), or tool comparison (SAKURA), is essential for confidence building.

4. Metrics and Safety Argumentation

A range of safety metrics is used, often cantered around Time-To-Collision (TTC), collision probabilities, and exposure-based residual risk. StreetWise goes furthest in formalizing residual safety risk by integrating exposure, failure likelihood, and injury severity across thousands of scenarios. SAKURA emphasizes remaining safety risk, benchmarking against a hypothetical attentive human driver.

In terms of safety arguments, only VVMethods clearly documents the use of structured approaches such as Goal Structuring Notation (GSN) and SysML-based modelling. Other frameworks acknowledge the need for formal documentation, though specific argumentation formats are often left undefined (e.g., EU Regulation, VMAD).

The survey and analysis conducted in Tasks T2.2 provided valuable insights into the diverse approaches taken across existing safety assurance frameworks. While a shared emphasis on scenario based validation was observed, differences emerged in terms of scenario generation methods, simulation strategies, safety argumentation formats, and the handling of uncertainty and in-service data. These findings have guided the development of the SUNRISE SAF by highlighting opportunities for improved consistency, interoperability, and methodological completeness. The resulting framework aims to complement and build upon prior efforts, offering a structured and harmonised approach aligned with international best practices and stakeholder expectations.

2.1.2 SAF alignment with Vehicle Safety Bodies

2.1.2.1 Alignment with UNECE

SUNRISE closely follows the regulatory activities of the United Nations Economic Commission for Europe (UNECE), particularly within the World Forum for Harmonisation of Vehicle Regulations (WP.29). Building on the NATM (New Assessment/Test Method) approach, the project integrates UNECE regulations into its SAF to foster international interoperability and lower barriers to global vehicle deployment. Through targeted technical input and structured feedback loops, SUNRISE contributes to the refinement of regulatory instruments—such as the interpretation document for the new regulation developed by the Working Group on Automated Driving Systems (ADS). This effort supports the cross-border harmonisation of vehicle safety in the context of CCAM (Connected, Cooperative and Automated Mobility) systems. Close collaboration with the Informal Working Group on ADS ensures strong alignment between ongoing UNECE regulatory developments and the SUNRISE SAF.

2.1.2.2 Alignment with EC and Member States

The SUNRISE project maintains close alignment with the European Commission (EC) and individual Member States to ensure coherence with overarching EU safety regulations and policy goals. This includes integrating requirements from existing vehicle safety legislation and preparing for future regulatory developments concerning automated and connected vehicles. By actively engaging with national transport authorities and EC representatives, the project contributes to the evolution of a harmonised Safety Assurance Framework, ensuring consistency across EU borders and fostering public trust. During the SUNRISE project, collaboration was established with several type approval authorities and technical services, including RDW (Netherlands), KBA (Germany), VCA (United Kingdom), UTAC (France), and IDIADA (Spain). A key outcome of the workshops conducted with these stakeholders was the identification of the need for application guidelines to support the auditing of the Safety Assurance Framework (SAF) from the perspective of an authority. These guidelines have been developed and are included in this document (see chapter 7) as well as in the interactive SUNRISE Handbook. In close cooperation with RDW, a mock-up approval assessment was also created, demonstrating a real-world application of the SUNRISE SAF. This collaboration was presented at the SUNRISE Final Event, with further details available in Deliverable D8.1 [27].

2.1.2.3 Alignment with Consumer Testing Organisations

In collaboration with consumer testing organisations such as Euro NCAP, the SUNRISE project aligns its safety assessment methodologies with independent evaluation criteria that strongly influence public perception and market acceptance. By incorporating consumer-focused performance metrics and real-world testing scenarios, SUNRISE ensures that its Safety Assurance Framework not only meets regulatory requirements but also addresses the high expectations of end users. This approach bridges the gap between legislative standards and practical safety outcomes. Throughout the project, several SUNRISE consortium partners actively participated in Euro NCAP working groups, and all other global NCAP organisations were contacted to explore alignment opportunities. A comprehensive overview of the global NCAP alignment is provided in SUNRISE Deliverable D8.1 [27].

2.1.2.4 Alignment with Standardisation

SUNRISE actively engages with European and international standardisation bodies such as CEN, CENELEC, ISO, SAE, BSI, and ASAM to ensure that its Safety Assurance Framework (SAF) is compatible with evolving technical standards. The project contributes to the development and adoption of common testing procedures, data formats, and system architectures that underpin the safe deployment of automated driving functions. By fostering alignment with ongoing standardisation efforts, SUNRISE ensures that its outcomes are applicable across diverse technological and industrial ecosystems, supporting interoperability and innovation.

In practice, the project closely aligned with key standardisation initiatives at international, European, and national levels. SUNRISE contributed to ISO working groups developing scenario-based safety assurance (e.g., ISO 34503 [5], 34502 [28], 34504 [29]) and structured safety case representations (ISO TS 5083). It also adopted elements of ASAM standards—such as OpenODD [30], OpenSCENARIO [25], OpenDRIVE [26], and OpenLABEL [31]—for scenario modelling, environment representation, and simulation structure. Though not formally part of ASAM, SUNRISE partners participated in several technical groups and helped shape specifications like OSI [32]. Alignment with SAE activities, particularly through J3279 [33] and J3206 [34], further anchored the SAF in accepted validation practices.

At the national level, the project collaborated with BSI in revising PAS 1883 [35] and developing Flex 1891 [36] and BSI 1889 [37], strengthening links between UK-specific standards and the SUNRISE framework. Across all engagements, SUNRISE maintained consistency in terminology and taxonomy by harmonising its internal concepts with industry standards, thereby ensuring coherence from scenario tagging to simulation and safety argumentation.

Overall, SUNRISE's SAF was well received by standardisation stakeholders for its modular, scenario-driven approach. Its technical architecture and simulation workflows were seen as aligned with key standards under development and already adopted across the industry.

2.1.3 SAF alignment with other EU projects

2.1.3.1 SELFY

The SELFY [38] project addressed **continuous assessments of the robustness and resilience** of CCAM-enabled mobility solutions against **cyber-attacks, malfunction, misuse or system failure** of the systems in use. It has built a set of tools to improve CCAM resilience and guarantee data security and privacy when different data is shared. SELFY's main goal was to promote a safe and secure operation amongst CCAM vehicles and mobility systems and services, enhancing trust and end-user adoption of CCAM solutions.

This section maps SELFY project components to the **Performance Assurance** pillar of the SUNRISE SAF, focusing on the **Scenario**, **Execute**, and **Safety Argument** blocks.

Scenario

- **Create**: SELFY enhances data-driven scenario creation through its SACP (Situational Awareness and Collaborative Perception) tool, which gathers real-time environmental data from vehicles and infrastructure. This enables detection of complex, high-risk situations such as near-misses or occluded hazards—ideal for generating realistic and safety-critical scenarios.

- **Format**: While SELFY doesn't directly handle formatting, outputs from SACP and CRHS (Cooperative Resilience and Healing System) provide semantic context and behavioural traces, supporting construction of logical and abstract scenarios.

- **Store**: Though not a scenario database, SELFY's TDMS (Trust Data Management System) ensures secure, traceable storage of scenario data, supporting integrity within a federated SAF ecosystem.

Execute

- **Concretize**: CRHS informs parameter selection for test scenarios by identifying conditions like sensor degradation or V2X issues, enhancing the strategic value of concrete test cases.

- **Allocate**: SELFY's SOTA (Secure Over-the-Air) updates ensure consistent, verified software across test environments, supporting reproducibility and execution integrity.

- **Execute**: During scenario execution, SACP provides live context monitoring while CRHS manages resilience responses. These tools capture real-time dynamics and system adaptations—e.g., fallback behaviour during comms failure—offering behavioural depth difficult to simulate otherwise.

Safety Argument

- **Coverage Analysis**: SACP provides contextual metadata (e.g., weather, traffic density) that refines ODD coverage evaluation. CRHS highlights under-tested conditions, feeding iterative test refinement.

- **Test Evaluation**: SELFY's VSOC (Vehicle Security Operations Center) monitors system-level anomalies, including safety-critical events and cybersecurity breaches. Even in collision-free runs, VSOC might flag internal faults or attack traces, enriching test verdicts with operational safety insights.

The mapping of SELFY components to the SUNRISE SAF demonstrates a complementary alignment between the two initiatives. SELFY's tools, such as SACP, CRHS, TDMS, SOTA, and VSOC, provide essential capabilities for identifying and assessing cyber-related anomalies. These components enrich the SAF's Performance Assurance pillar by strengthening its ability to capture/identify/assess complex interactions, resilience behaviour, and system vulnerabilities under cyber stress conditions. SUNRISE SAF users are presented with the feasibility to incorporate SELFY's cybersecurity assessment protocols seamlessly within SAF execution workflows, particularly when addressing safety impacts stemming from cyber threats.

2.1.3.2 i4Driving

Another EU collaboration that took place within the SUNRISE project is with the i4Driving project [39], whose objective is to create naturalistic human driver model for the development and assessment of CCAM systems. Various driving simulator experiments were carried out across different sites, the data collected were then used to create the i4Driving human driver model. Specifically within the i4Driving project, work package 6 then takes the human driver model, applies common and harmonised safety evaluation method to evaluate it against the use cases. Due to the timewise overlap of SUNRISE and i4Driving (after the 1st year of the i4Driving project), an earlier draft version of the SUNRISE SAF has been adopted by the i4Driving project as their evaluation workflow. This includes all the performance assurance components, covering scenario creation, formatting, database storage and query, simulation environment allocation, virtual test execution and safety analysis.

Within each of the five unique applications, workshops were conducted within the i4Driving project to walk through the whole SUNRISE SAF, starting from the scenario block. UoW has acted both as the i4Driving work package lead on evaluation, and the SUNRISE T2.2 lead on SAF, forming a seamless integration and collaboration between the two projects. In addition, i4Driving project member RDW has been heavily involved in the SUNRISE expert platform, stakeholder engagement, and mock up type approval exercise.

This collaboration has resulted in a close alignment between the SUNRISE SAF and the i4Driving human driver reference models. The SAF's performance assurance workflow was directly applied in i4Driving's evaluation work package to benchmark naturalistic human driving behaviour against relevant CCAM use cases. Through joint workshops and overlapping leadership, a bidirectional exchange was established: the i4Driving project adopted the SAF as its methodological backbone for human driver model evaluation, while feedback from the i4Driving use cases complimented the SAF's components. SUNRISE SAF users can therefore build upon this connection by adopting i4Driving human driver models within the Test Evaluation block to benchmark CCAM system performance, or within the Execute block to form a combined SUT with human driver and CCAM system.

2.1.4 SAF alignment with the Expert Platform

This section explains the cooperation with the SUNRISE Expert Platform, and how the SUNRISE SAF has been aligned to the feedback obtained from its members.

The Expert Platform is a group of external individual CCAM stakeholders, through which the SUNRISE project gathers feedback on its plans and results. Subsequently, that feedback is taken into account during the remaining course of the project and the development of the SAF and its adoption by target users.

As can be observed in the following figure, the Expert Platform is subdivided into the Industry Board (IB) and the Advisory Board (AB). This figure shows that the Expert Platform consists of both stakeholders from industry entities like OEMs and TIER1s, *and* non-industry entities like standardization bodies, regulatory authorities and policy makers.

Figure 5: Expert Platform (EP) = Industry Board (IB) + Advisory Board (AB)

Expert Platform candidates were invited through several rounds of invitations. Becoming a member of the Expert Platform, requires a subscription through the SUNRISE website through the following link:

https://ccam-sunrise-project.eu/register/

The Expert Platform was established in July 2023 and the total number of registered members reached around 150.

The SUNRISE consortium cooperated with the Expert Platform in 2 different ways: [1] through the Cooperation Platform and [2] through online workshops, each of which will be treated in further details in the following sections.

2.1.4.1 Cooperation platform

The Cooperation Platform is a discussion forum integrated into the SUNRISE website, with the purpose of providing a continuous communication channel between members of the SUNRISE consortium and the Expert Platform. This forum allows members of the SUNRISE consortium to launch discussions, interact and cooperate with the members of the Expert Platform. The inputs registered on the Cooperation Platform, are used within the SUNRISE project to improve the development of its Safety Assurance Framework and other project results.

At the start of the SUNRISE project, IDIADA, ERTICO and TNO defined a set of requirements for the Cooperation Platform, after which it was integrated into the SUNRISE website. It was brought into operation in February 2023. And by July 2023, the total number of registered members reached around 150. After that, feedback on SUNRISE plans and results has been obtained from Expert Platform members by means of their written comments on numerous questions published on the Cooperation Platform. These comments were mainly provided during a series of Expert Platform workshops, which will be explained in the next section. The following figure shows some example questions and answers on the Cooperation Platform.

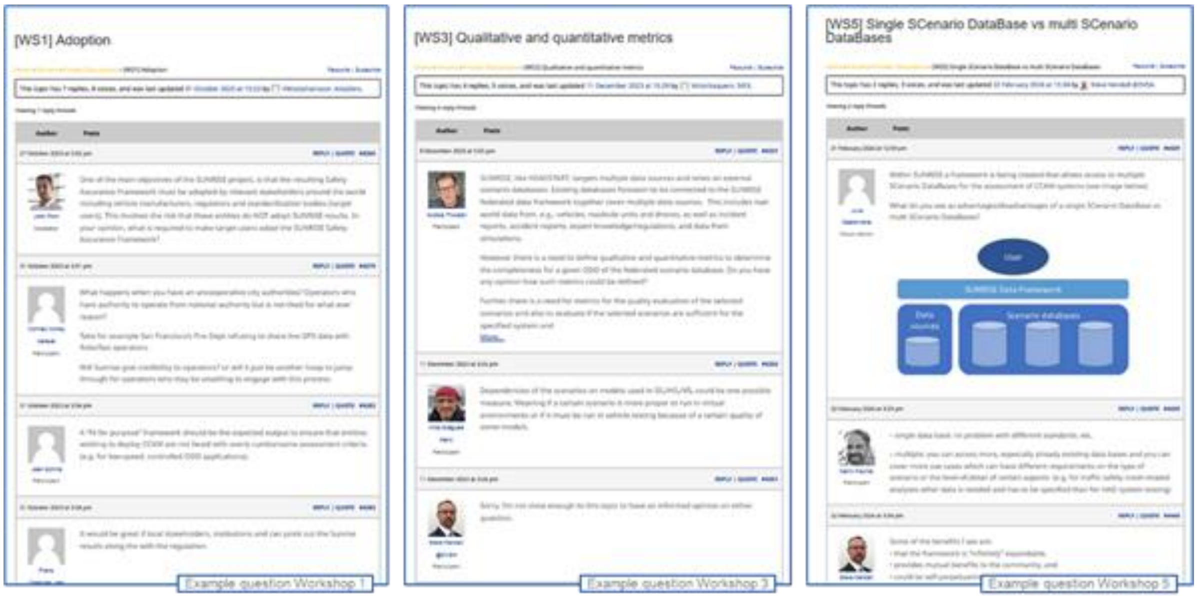

Figure 6: Example questions and answers on the Cooperation Platform

Once registered, the Cooperation Platform can be accessed through the following link:

https://ccam-sunrise-project.eu/forums/

2.1.4.2 Workshops

A total of 10 workshops have been celebrated with members of the SUNRISE Expert Platform. These workshops were organized and celebrated to inform Expert Platform members about project plans and results, and (most important) to obtain their feedback on it.

Figure 7: Some of the attendees to some of the Expert Platform workshops

The table below provides an overview of all workshops including their topics (*Topics*), the people involved (*By*) and when the workshop was celebrated (*When*). As can be seen in this table, workshops 1-6 mostly focussed on the general explanation of the project, its work packages and the first deliverables produced, whereas workshops 7-10 focussed on external feedback resulting from the SUNRISE Midterm Event (celebrated on 24th of April 2024 in Brussels), where the first draft version of the SAF was presented.

Table 1: Overview Expert Platform Workshops

Nr.	Topics	By	When
1	A. SUNRISE Introduction B. Work package description C. Project goals	Leaders of WP1, T9.4	OCT 2023
2	A. WP2 Description B. D2.1 Summary - Overview and gaps of existing safety assessment frameworks	Leaders of WP1, WP2, D2.1	NOV 2023
3	A. WP3 Description B. D3.1 Summary - Report on baseline analysis of existing Methodology	Leaders of WP1, WP3, D3.1	DEC 2023
4	A. WP4 Description B. D4.1 Summary - Report on relevant subsystems to validate CCAM systems	Leaders of WP1, WP4, D4.1	JAN 2024
5	A. WP5 Description B. D5.1 Summary - Requirement for CCAM safety assessment data framework content	Leaders of WP1, WP5, D5.1	FEB 2024
6	A. WP7 Description B. D7.1 Summary - CCAM Use cases validation requirements	Leaders of WP1, WP7, D7.1	MAR 2024
7	***Draft* Safety Assurance Framework** Explanation and feedback on a draft version of the SAF in a dedicated session during the SUNRISE Midterm Event in Brussels, to which Expert Platform members participated online.	Leaders of WP2, T2.1	APR 2024
8	**Industry engagement** Goals: Increase focus on SAF users in the industry (OEMs and TIER1s). Update them on draft SAF, including both the framework and its handbook. Collect and process their requirements, aiming for their engagement and adoption of the SAF. Convince and encourage them to start using the SUNRISE SAF, thereby boosting its uptake and after-life.	Leaders of WP1, WP2, T2.2, WP8	NOV 2024
9	**Role of society** Goals: Increase focus on role of wider public or society, ensuring to gain their trust and meet their safety expectations, while dully informing about the limitations of CCAM systems and their safety. To help achieving this, cooperate with EU-funded projects like SINFONICA [40] and Cultural Road, and CCAM safety related associations like MOVING and PAVE [41].	Leaders of WP1, WP2, WP9	JAN 2024

10	**Alignment certifiers & industry** Goals: Increase focus on alignment between certifiers and industry regarding CCAM safety assurance. Aim for a SAF that is used by both groups in the same or similar way. Aim for a SAF that covers shared needs and expectations to the highest possible extend. Identify differences in SAF use in both groups, and make sure to cover these adequately.	Leaders of WP1, WP2, T2.2, WP8	FEB 2024

The workshops have benefitted and enriched the SUNRISE project and its SAF in various ways, thereby aligning the SAF with the feedback provided by the members of the Expert Platform. The main benefits and enrichments obtained in the workshops, are explained hereafter.

During **workshops 1-6** feedback was obtained by means of written comments on numerous questions published on the Cooperation Platform. In total 25 questions were raised, to which 74 written answers were provided by Expert Platform members. The exact questions and answers can be found on the Cooperation Platform (through the link provided earlier). The main **conclusions** (in order of priority) drawn from all feedback received through the Cooperation Platform are:

1. **Definition of the right metrics and KPIs** and their monitoring seems to be commonly considered of importance.
2. **Making the SAF future proof** is found to be of importance. That involves predicting the (foreseeable) future, and updating various SAF items over time (like scenarios, metrics and KPIs).
3. **SAF application in CCAM pilot projects** on public roads (or similar real-life applications), would probably benefit the project and the adoption of its results.
4. Confirmation that the **role of standards is paramount**. This applies for example to scenarios, database contents and interfaces. But also, to the handling of new or conflicting standards.
5. Mechanisms for the **identification of unknown unsafe scenarios**, seems to be only limitedly covered by the current project plans and results.
6. The **public** or **society** should probably play a **bigger role** in the project. For example to gain their trust and ensure that their safety expectations are met.
7. More focus on the **role of the human driver** with regards to safety assurance of CCAM systems. Their unpredictable behaviour is likely to significantly impact the safety of CCAM systems.
8. More focus seems to be required on the **alignment between certifiers and industry** regarding the safety assurance of CCAM systems.

In **workshop 7**, attendants were provided with access to an online draft version of the Safety Assurance Framework (on Miro.com), on which they could provide their feedback in the form of digital “sticky notes” (see figure below). These comments were then taken into account into the further development of the SAF.

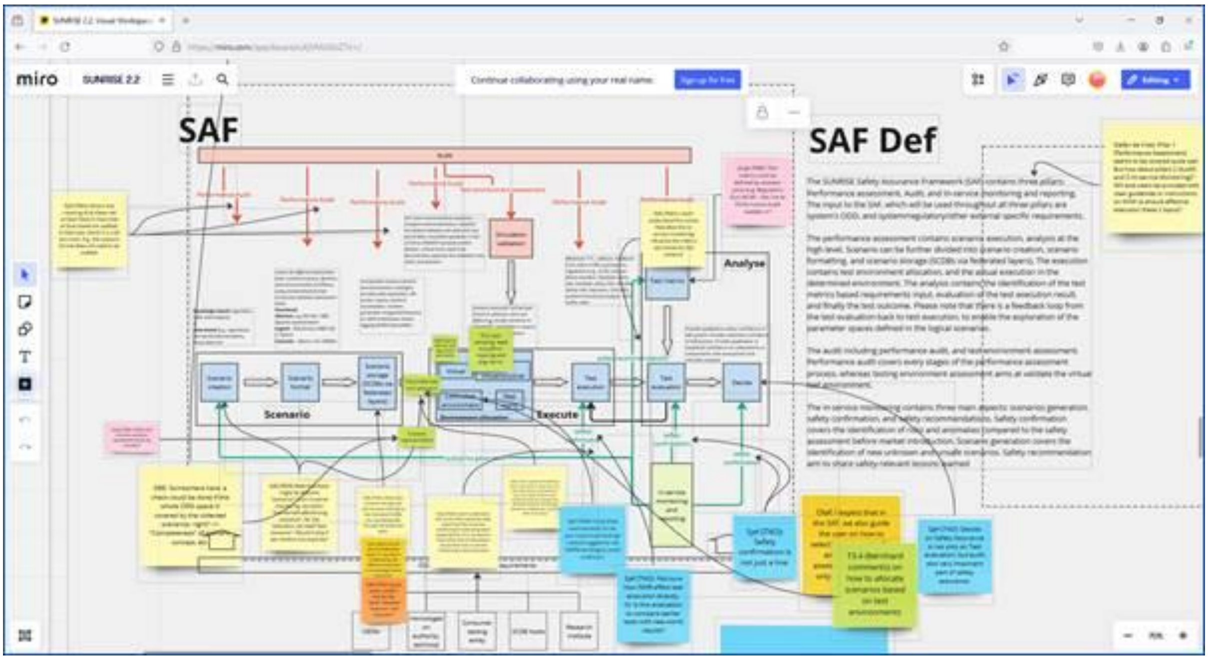

Figure 8: Example of feedback on SAF through “sticky notes” on Miro.com

During **workshop 8** potential SAF users from the industry (among which employees of Mercedes Benz, BMW and Volvo) were updated on the current version of the SAF. They also provided (verbal) answers on a series of predefined questions. Their feedback has been processed by the task and work package leaders involved in this workshop.

In **workshop 9** SUNRISE invited representatives of the SINFONICA [40] and OPTIPEX [42] projects. Each project was shortly introduced, after which an open discussion was sparked by a few predefined questions related to the role of society in safety assurance of CCAM systems. The main **conclusions** following from that discussion were as follows:

1. **Main take-away**: Address general public or specific CCAM user groups (like disabled, elderly, user associations), highlighting CCAM benefits and transparently communicating CCAM safety related matters. This will help to gain trust and manage expectations.
2. Both SINFONICA and OptiPex partly focus on **personal safety**, and only very limitedly on vehicle safety.
3. CCAM systems are generally **trusted to be safe** by the general public.
4. **Possible connections** to the SUNRISE SAF through the ISMR block (in-service monitoring and reporting), and by possibly taking public trust issues into account in the evaluation criteria, and by ensuring the existence of scenarios that are closely related to public trust, such as scenarios involving pedestrians or disabled people.

These conclusions have been addressed by the task and work package leaders involved in this workshop. The main take-away (conclusion number 1), is being addressed by publishing a series of frequently asked questions (FAQ) and answers on the SUNRISE website, in cooperation with the CONNECT [43] project (due August 2025). See Section 3.7 of deliverable D9.3 [44] for further details on this topic.

In **workshop 10** representatives of both certifying and industry entities were invited, aiming for a SAF that is used by both groups in the same or similar way. After introducing the SUNRISE project and its SAF, an open discussion was sparked by various predefined questions related to alignment of the SAF with the needs and expectations of both groups. The main **conclusions** following from that discussion were as follows:

1. **Different roles in applying the SAF**
   Vehicle Safety Bodies emphasized that, unlike industry users, they typically do not perform all steps of the SAF themselves. Instead, their role is more focused on auditing or witnessing the SAF activities carried out by others, rather than executing the process end-to-end.
2. **Need for guidance from a perspective of a Vehicle Safety Body**
   Vehicle Safety Bodies once again highlighted the need for clear guidance on how to apply the SAF from an authority or certification perspective. There is a strong demand for structured documentation and practical instructions tailored to the specific needs of regulatory and certifying bodies.
3. **SUNRISE as a driver for advancing CCAM safety assurance**
   Both Vehicle Safety Bodies and industry stakeholders agreed that the SUNRISE project could provide valuable methods and application examples that support the further development of safety assurance practices for CCAM systems. The project's outcomes can be seen as a meaningful contribution to aligning industry and regulatory efforts.

2.1.5 SAF alignment with international twins

This section explains the cooperation with twin projects and initiatives, and how the SAF has been aligned to the feedback obtained from them.

Through various online meetings and international events in the first half of the project, strong connections have been established and collaboration talks have been initiated with key representatives of international twin projects and initiatives. During the further course of the project, these strong connections have been exploited in various ways, with the most relevant ones explained in the following sections.

In this context, projects and initiatives are indicated as "twins" if cooperation went beyond making use of their results, and involved bi-directional interactions such as (online) meetings and joint participation on events or other activities. Furthermore and without downplaying their importance, twin projects and initiatives do *not* include other EU-funded projects in this context. The most notable twin projects and initiatives that comply with this are: V&V Methods , Sakura [11], AVSC [45], KATRI [46] and Transport Canada [47].

The SUNRISE project also made use of the results and knowledge of many other non-EU-funded projects and initiatives. However, without bi-directional interactions as mentioned above. The most notable examples of this type of projects and initiatives are: SET Level [48] and DIVP [49] (both of which are closely related to the V&V Methods and Sakura projects).

2.1.5.1 Benchmark

In many occasions, SUNRISE team members used the results or approaches applied by twin projects and initiatives as a reference or benchmark in the creation of SUNRISE deliverables. That often involved previous discussions between experts on both sides, and lead to improved alignment between the SUNRISE SAF and other relevant safety assurance frameworks and approaches.

The most notable examples where twin projects and initiatives have been used as a reference or benchmark for SUNRISE deliverables, are summarized in the following table.

Table 2: Examples of SUNRISE deliverables influenced by twin projects and initiatives

D2.1_Overview-and-gaps-of-existing-safety-assessment-frameworks_V1.1	
Twins	V&V Methods project (under Pegasus), Sakura project
Influence	The CCAM safety assurance approaches applied by the indicated twins, have been used as a benchmark for the creation of the SUNRISE SAF, in which both their strengths and weaknesses (gaps) have been analysed and addressed.
D3.1_Report-on-baseline-analysis-of-existing-Methodology_V1.1	
Twins	V&V Methods project (under Pegasus), Sakura project
Influence	The scenario-based safety assurance methodologies applied by the indicated twins, have been used as a benchmark for the creation of the SUNRISE Methodology. For that, the HEADSTART methodology is taken as a baseline and complemented with existing best practices like those of the indicated twins.
D5.1_Requirement-for-CCAM-safety-assessment-data-framework-content_V1.1	
Twins	V&V Methods project (under Pegasus), Sakura project
Influence	The way the indicated twins apply scenarios (and related data) for safety assurance of CCAM, has been used as a reference for the definition of requirements for the SUNRISE Data Framework and the external scenario databases that connect to it.

As can be observed in the table above, various twin projects and initiatives formed an important reference or benchmark for laying the foundations of the SUNRISE Safety Assurance Framework.

2.1.5.2 Feedback at Midterm Event

At the SUNRISE midterm event (celebrated on April 24th in Brussels), a draft version of the SUNRISE SAF has been presented. Key representatives of twin projects and initiatives joined that event, and provided their feedback on the draft version of the SAF (on which they were previously briefed). Representatives of V&V Methods, Sakura, AVSC and KATRI joined that event in person. Transport Canada joined online.

Figure 9: Twin projects and initiatives providing feedback on draft SAF at SUNRISE Midterm Event

The feedback received on the draft SAF has been of great value for its further refinement, during the 2nd half of the SUNRISE project. For that reason, it has been carefully documented in an Excel spreadsheet, in which specific feedback has been assigned a priority level (in green, yellow or red) and a responsible SUNRISE partner, that defined an adjustment plan. During the further course of the SUNRISE project, regular follow-ups on each feedback item were then undertaken to ensure progress on each adjustment plan. In that way, all feedback items were addressed and closed towards the end of the project. A few examples of feedback items on that Excel spreadsheet can be observed in the following figure.

Figure 10: Snippet of Excel with external feedback on draft SAF

The midterm feedback received on the draft SAF, lead to the most significant alignment of the SUNRISE SAF with the safety assurance frameworks and approaches of twin projects and initiatives.

2.1.5.3 Feedback at Final Event

Representatives of various twin projects and initiatives attended the SUNRISE Final Event at IDIADA headquarters in Santa Oliva (Spain) on June 18th 2025. During that event, consortium-external attendees provided their feedback on a series of predefined questions about the SUNRISE project and its SAF. Among these respondents were representatives of the following twin projects and initiatives: V&V Methods, Sakura and AVSC.

2.2 Policies, standards, regulation and legislation

This section explains the close connections of the SUNRISE SAF with policies, standards, regulations and legislation. On one hand, the SUNRISE consortium undertook efforts to achieve the possible inclusion of (parts of) the SUNRISE SAF in policies, standards, regulation or legislation. On the other hand, the development of the SUNRISE SAF has been influenced by a large number of standards and regulations. Further details of these connections are explained hereafter.

2.2.1 Influence on

The SUNRISE project has developed a practical CCAM Safety Assurance Framework (SAF). Among the target users of this framework are vehicle safety regulators and standardization bodies. SUNRISE's Work Package 8 cooperates with these target users, aiming for their engagement and adoption of the SAF. This includes the influencing and possible inclusion of (parts of) the SUNRISE SAF in policies, standards, regulation or legislation.

At the moment of writing this report, it is too early to clearly indicate the impact that the SUNRISE project will have on specific policies, standards, regulation or legislation. However, at this stage certain tendencies can be detected (both inside and outside the domain of the European Commission), as explained hereafter.

Since the Technology Readiness Level (TRL) of the SUNRISE SAF has been raised to 7, further maturing will be needed. Therefore, the final results of the SUNRISE project are highly likely to include recommendations of the drafting of new Horizon Europe calls by the European Commission (EC).

In 2022, the EC adopted the first worldwide legislation concerning the type-approval of the Automated Driving Systems of fully Automated Vehicles: EU Regulation 2022/1426. To ensure the establishment of harmonized practices across the EU, the EC launched the process of drafting a first Interpretation Document. More recently, the EC's Joint Research Center (JRC) has started working on the creation of a second Interpretation Document. Through task T8.2, SUNRISE aims to align its main findings and results with this second Interpretation Document.

Beyond the domain of the EC, and through task T8.1, SUNRISE pursues the engagement and cooperation with UNECE's Working Party on Automated/Autonomous and Connected

Vehicles: the GRVA. The GRVA supervises a series of informal work groups (IWGs) and tasks forces. One of those IWGs is on ADSs, and has received a mandate to draft a UN Regulation and a Global Technical Regulation (GTR) on automated driving. The GTR focuses on technical requirements, while the UN Regulation combines the regulatory text with specific administrative provisions. Through task T8.1, SUNRISE aims to align its main findings and results with these regulations created by the ADS IWG.

Also beyond the domain of the EC, through task T8.3 and other channels, SUNRISE pursues the engagement and cooperation with consumer testing entities, mainly focusing on Euro NCAP. Various SUNRISE partners participate in Euro NCAP working groups and use SUNRISE findings and results to further develop the Euro NCAP protocols. The other way around, Euro NCAP related experts influence the SUNRISE results by communications through the Expert Platform or via direct feedback to SUNRISE partners. So through task T8.3 (and other channels), SUNRISE aims to align its main findings and results with the future protocols created by Euro NCAP.

2.2.2 Influenced by

The SUNRISE project is influenced by several existing standards and regulations. What follows are the main standards and their usage within the project:

Foundation Standards

1. **UNECE's NATM (New Assessment/Test Method for Automated Driving)**
   A. Forms the foundational structure for SUNRISE's SAF
   B. Provides basis for multi-pillar approach in safety assessment
2. **ISO 34501:2022 - "Road Vehicles - Test Scenarios for Automated Driving Systems - Vocabulary"**
   A. Guides the establishment of scenario-based terminology aligned with ISO 21448 and ISO 26262
   B. Provides fundamental terminology for scenarios and testscenario concepts
   C. Establishes framework for scenario descriptions (abstract, functional, logical, concrete)
3. **ISO 34502:2022 - "Scenario-based safety evaluation framework"**
   A. Guides definition of scenario-based safety evaluation framework
   B. Guides identification and risk evaluation of hazardous scenarios of the ADS
   C. Guides validation of test methods and simulation quality
   D. Provides criteria for test platform allocation
4. **ISO 34503:2023 - "Road Vehicles - Test Scenarios for Automated Driving Systems - Taxonomy for ODD"**
   A. Provides base structure for scenario comparison and classification
   B. Provides taxonomy for ODD and environment conditions
   C. Establishes framework for dynamic elements classification
5. **ISO 34504 - "Scenario categorization" (under development)**
   A. Guides on definition of scenario categorization approach
   B. Provides structure for test cases
   C. Guides on inclusion of tags within scenario definition

Data Exchange and Testing Standards

6. **ASAM OpenSCENARIO** [25] **and OpenDRIVE** [26]
   A. Scenario formatting and description
   B. Used in multiple simulation tools
   C. Defines behaviour of dynamic actors
   D. Road network description
   E. Essential for automated execution in simulation environments
   F. Referenced for scenario description metrics
7. **ASAM OpenLABEL** [31]
   A. Guides on scenario tagging and metadata
   B. Provides comprehensive mechanism for tag construction
   C. Guidance on multi-sensor labelling
   D. Used for standardized annotation formats and labelling methods
   E. Guidance on object-level information representation
8. **ASAM OSI (Open Simulation Interface)** [32]
   A. Guidance on test execution results description
   B. Provides standardized interfaces for simulation models
   C. Guidance on sensor data exchange
   D. Guidance on vehicle dynamics communication
9. **ASAM OpenTestSpecification** [50]
   A. Guidance on test evaluation and criteria definition
10. **ASAM OpenODD (under development)**
    A. Guidance on ODD taxonomy standardization
    B. Provides a machine-interpretable format for ODD specification
11. **FMI (Functional Mock-up Interface)** [51]
    A. Provides information on model exchange and integration across different simulation tools
    B. Provides co-simulation support
12. **SAE J3016** [52]
    A. Used for general scenario database metrics
    B. Possible consideration of DDT terminology
    C. Provides framework for scenario classification
    D. Provides taxonomy for driving automation systems

Safety Standards

13. **BSI PAS 1883:2020 - "ODD taxonomy for ADS"** [35]
    A. Provides foundation for ODD taxonomy
    B. Guidance on integration with ASAM OpenLABEL ontology
14. **BSI Flex 1889:2024 – "Natural Language Description of scenarios for ADS"** [37]
    A. Provides structured natural language format for Level 3+ ADS test scenarios
    B. Aids in creating, categorizing, and communicating test scenarios
15. **ISO 26262 (Functional Safety - FuSa)** [2] **and ISO 21448 (SOTIF)** [3]
    A. Guides safety-related requirements
    B. Provides a reference framework for safety assessment
    C. Guidance on unreasonable risks from functional deficiencies
    D. Guidance on analysis of safety in use and external factors

E. Guidance on evaluation of unknown-unsafe scenarios
F. Guidance on assessment of system behaviour in known unsafe scenarios
G. Provides information on testing at ODD boundaries
H. Used as reference for scenario database quality metrics

16. **UNECE R157 (ALKS Regulation)** [53]
    A. Used as reference for driver models in critical situations in the context of the ALKS function
    B. Provides requirements for automated driving systems validation
    C. Defines criteria for "reasonably foreseeable" scenarios in the context of the ALKS function
    D. Used for scenario exposure metrics

2.3 Scenarios vs test scenarios

This section aims to clarify the conceptual distinction between a *scenario* and a *test scenario*, which is essential for understanding how safety validation is structured within scenario-based approaches. While both terms are often used interchangeably, they serve different purposes: scenarios describe what can happen, whereas test scenarios are specific inputs for testing system behaviour under defined conditions. Recognising this distinction is critical for ensuring traceability, relevance, and completeness in safety assessments.

2.3.1 Scenario

In various databases, scenarios are collected that are used to describe what a CCAM system may possibly encounter in operation during its lifetime (or deployment time). The scenarios in these databases provide a structured view on what has happened on the road or what may possibly happen. Scenario databases are set up for various purposes, not only for safety assessment [54].

An accident database provides a detailed view on what accidents happened on the road, what were the scenario(s) leading to the accident, and what conditions applied. An accident database also provides an indication of the severity of the injuries of the people in the accident. Other scenario databases identify, characterize and record all scenarios from vehicles driving on the road – usually a limited Number of vehicles driving for many kilometres (comparable to Field Operational Tests, or Naturalistic Driving Studies). Such scenario databases are mainly data-driven. Alternatively, scenarios can be described based on knowledge and experiences. Also such knowledge-based scenarios can be added to a scenario database.

A scenario describes any situation on the road including the intent of the ego vehicle, the behaviour of road users, the road layout, and conditions such as weather and lighting. A drive on the road is considered a continuous sequence of scenarios – which might overlap. More formally according to [55], a scenario is a quantitative description of the relevant characteristics and activities and/or goals of the ego vehicle(s), the static environment, the dynamic environment, and all the events that are relevant to the ego vehicle(s) within the time interval between the first and the last relevant event. An event corresponds to a moment in time at which a mode transition occurs or a system reaches a specific threshold, where the former can be induced by both internal and external causes. A shorter version of the definition of scenario is also mentioned within the SUNRISE project as the ‘description of a temporal and spatial traffic constellation’.

2.3.2 Test scenario

Test scenarios form the input to scenario-based safety assessment such as virtual testing or track testing. Following the IEEE definition of a test, which is clearly explained in [56], a test is an evaluation of: *a statement on the system-under-test (test criteria; what is evaluated using the test), under a set of specified conditions (test scenario; how are the test criteria evaluated), using quantitative measures (metrics; how is the outcome of the test expressed quantitatively), and a reference of what would be the acceptable outcome (reference; when is the outcome acceptable).* Note: the reference may be formulated for each individual test, but also for a set of tests, e.g. in case of risk quantification.

Hence, a test scenario may not be identical to a scenario in a scenario database. Because of sampling (and interpolation), it might be a case that has never been seen exactly like that in the real world. A test scenario might also not contain all information that is available from the scenario database, and a test scenario might be complemented with models or other information to be able to correctly perform a test. Additionally, a test scenario might be a simplification of a (real-world) scenario. A test scenario of Euro NCAP to test emergency functions such as AEB, is usually defined with one vehicle and one collision partner. The real-world scenario on which such test scenario is based, may easily have contained more than one other road participant, and a different static environment than the environment represented on the test track. Assumptions are made in the translation from (real-world) scenarios to test scenarios. The audit, should assess how valid these assumptions are in the selection and generation of test scenarios.

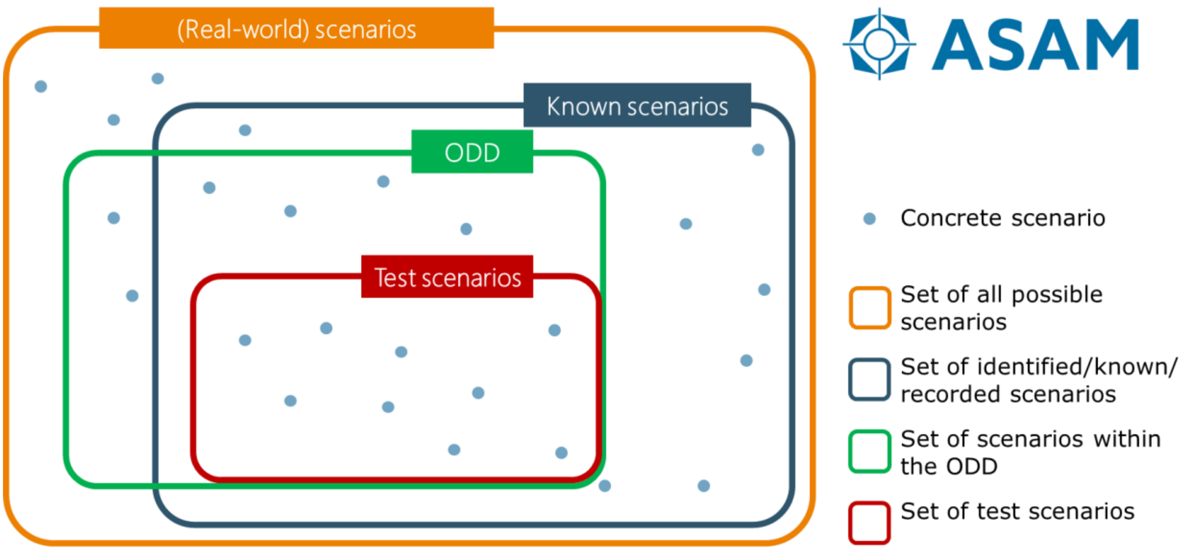

Figure 11: A schematic view on the relation between scenarios, the operational design domain, and test scenarios. Adapted from [57].

Figure 11 presents a schematic view on the relation between the scenarios in the real world, the known scenarios that are collected in the scenario databases, the scenarios in the ODD and the scenarios used to generate the test scenarios [57]. With this scheme, the important concepts of completeness and coverage can be understood:

- Completeness: how well do the known scenarios cover all possible scenarios in the real-world? How to estimate completeness is shown in [58];
- Coverage: how well do the selected test scenarios cover the ODD of the system under test? Ideally, there are no unknown scenarios in the ODD and the set of test scenarios covers at least the complete ODD.

The figure above illustrates the importance of striving for completeness in a scenario database. Not only is it difficult to provide a reliable description of the ODD of a function when the scenario database is not sufficiently complete, also the relevance of the selection of test scenarios is limited in that case. In other words, the function might encounter a scenario in reality for which the function has not been tested.

3 INPUTS TO THE SAF

Figure 12 highlights the Input block of the SAF.

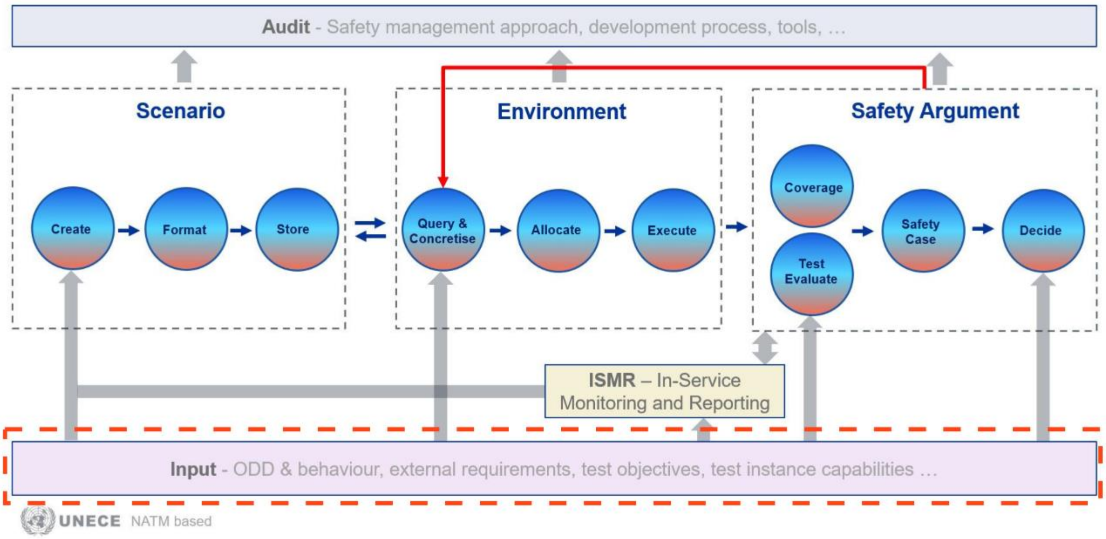

Figure 12 : Input block of the SAF

There are five arrows that input into other SAF blocks originating from the **Input** block. The information covered by the Input block includes items like description of the Operational Design Domain (ODD), external requirements, CCAM system specifications, variables to be measured during test execution, pass/fail criteria for successful test execution, and monitoring requirements.

At all five arrows, the **ODD** description of the CCAM system is considered. The ODD description includes ranges of relevant attributes, such as the maximum vehicle speed and rainfall intensity it is able to handle. The ODD description may adopt an inclusive approach (describing what is inside the ODD), an exclusive approach (describing what is outside the ODD), or a combination of the two. It would be highly beneficial if the ODD description is standardized and machine readable, the SUNRISE Ontology (described in deliverable D5.2 [59]) proposes an ODD (ISO34503) [5] and behaviour (ASAM OpenLABEL [31]) aligned ontology. For formatting the ODD, it is suggested to follow the norms related to the ODD definition format listed in ISO 34503 (2023) [5].

The Create block can utilize the ODD description to create scenarios that are part of the described ODD.

The Query & Concretize block uses the ODD description to generate the test cases that are needed for the safety assurance of the CCAM system, including the description of the needed output to analyse the results.

The Test Evaluate block uses the ODD description to determine whether the CCAM system operated safely within its ODD in a specific test.

The Coverage block uses the ODD description to check whether the ODD space is sufficiently covered.

The ISMR block (in-service monitoring and reporting) employs the ODD description in order to verify whether the system – once it is deployed on the public road after certification by the vehicle authority – is operating inside its ODD.

Besides the ODD description, various blocks also need **the external requirements** applying to the CCAM system. These requirements should reflect the required behavioural competences, regulations, rules of the road, safety objectives, standards and best practices. The requirements can be a source for creating scenarios, which is why the requirements are part of the first interface. Furthermore, it is important that the Query & Concretize block considers the requirements and outputs relevant test cases, as the goal of the SAF is to assure that the requirements are met.

Note that this process also establishes the means to measure compliance with the requirements for the test cases. Not all requirements can be formulated using test validation criteria, which is why the requirements can also be communicated to the Test Evaluate block.

Lastly, the input to the ISMR block needs requirements to check whether system-level requirements are satisfied over the lifetime of the system. Note that requirements can be very different from system to system, so formalizing this might be challenging. For that reason, a standardized description format would be preferred.

The **CCAM system** is the main subject of the test cases, thus its technical specifications need to be provided to the Query & Concretize block. The CCAM system can also be a source for creating scenarios. For example, scenarios created using knowledge of the system architecture and fault analysis techniques such as systems-theoretic process analysis (also known as STPA [60]). The CCAM system can be a physical prototype or virtual model of the actual system (or a combination of both).

In case some variables need to be measured during the test execution – additional to the variables that are needed to verify the test objectives – this information can be provided to the Query & Concretize block.

Additionally, pass/fail criteria for successful test execution have to be provided to the Query & Concretize block, in order to identify when test cases have or have not successfully been executed. For example, if there are certain tolerances on speed values or lateral path deviations, these shall be included. When the test case is executed outside of the pass/fail criteria, then its execution would have to be deemed as unsuccessful from an execution point of view. These examples are related to both real world and test track test allocation, but other examples may apply to virtual test environment (for example when simulation output contains an error).

For all inputs, it is assumed that simulation models (other than those representing the CCAM system) and the simulation platform are part of the “Execute” component and, therefore, not provided externally. Hence, simulation models (other than that of the CCAM system) and the simulation platforms are not part of the listed interfaces.

4 PERFORMANCE ASSURANCE

Figure 13 highlights the SUNRISE SAF with expanded Performance Assurance pillar.

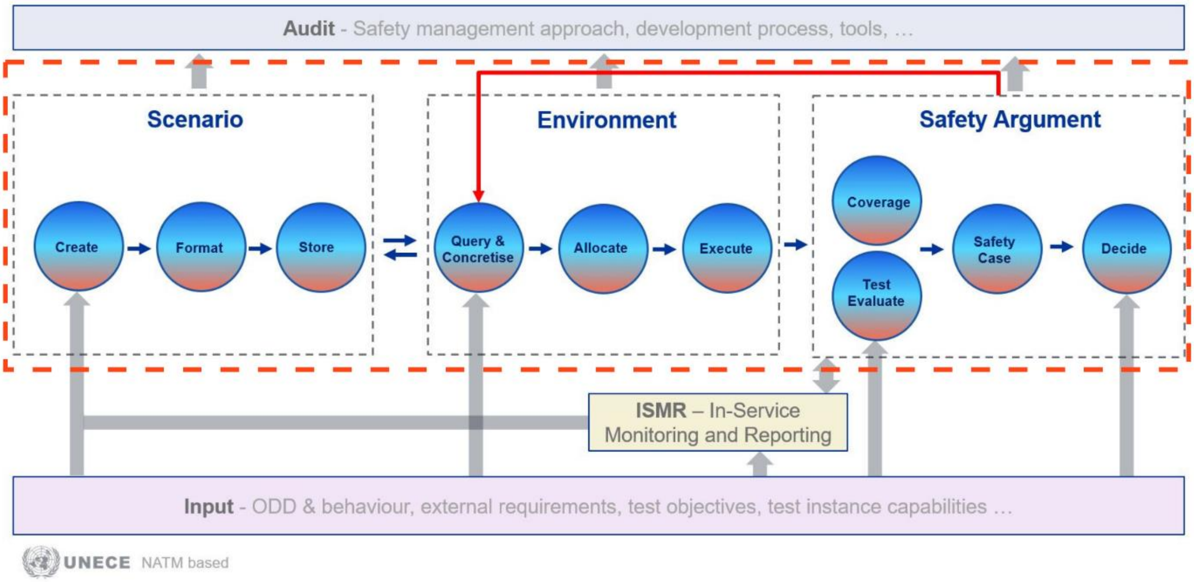

Figure 13: SUNRISE SAF with expanded Performance Assurance pillar

The SUNRISE SAF is a harmonized structure of methods, tools and data that allows stakeholders to determine whether a CCAM system meets a set level of safety, for introduction and during deployment.

The SAF is based on the multi-pillar approach from UNECE's NATM [1].

The **Safety Performance Assurance** block includes but is not limited to methods and tools that:

- create, format, store and extract relevant scenarios
- allocate relevant scenarios to different test environments
- execute relevant scenarios in different test environments
- analyse and evaluate test results
- elaborate evidence-based safety argumentation and decisions

In the SUNRISE project, the focus is on Performance Assurance using a scenario-based approach and incorporating virtual, hybrid- and physical test environments. In addition, it features the SUNRISE Data Framework which allows users to obtain scenarios from multiple qualified scenario databases (see corresponding section for details).

At a high level, the Safety Performance Assurance block comprises the Scenario, Environment, and Analyse sub-blocks. Essentially, this means that to conduct the safety assurance process, a set of scenarios is needed, an environment for executing these scenarios, and a process to analyse the execution results to determine whether the system is safe.

4.1 Scenario

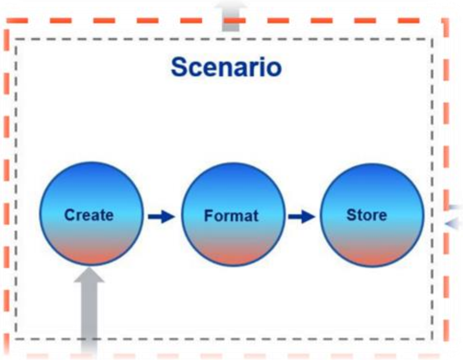

Figure 14 : Scenario block of the SUNRISE SAF

The **Scenario** block focuses on managing scenarios required for safety assurance, encompassing three key processes: creation, formatting, and storage, see Figure 14.

1. **Create** involves generating scenarios based on data-driven or knowledge-driven approaches. Scenarios are derived using methodologies like StreetWise [12], which extracts real-world scenarios and statistics from driving data, or a database like Safety Pool™ [13], which leverage various scenario generation methods. These processes occur within individual SCDBs, reflecting database-specific requirements and use cases. While scenario creation is handled independently by SCDB providers, it is integral to the overarching safety assurance framework.
2. **Format** structures scenarios into appropriate representations for effective communication and downstream testing. Multiple levels of abstraction are employed to cater to diverse stakeholders, ranging from human-readable functional and abstract-level scenarios to machine-readable logical and concrete-level scenarios. Common formats like ASAM OpenSCENARIO [25], OpenDRIVE [26], and BSI Flex 1889 [37] ensure interoperability and standardization.

- **Store** consolidates formatted scenarios within SCDBs, enabling seamless access and integration. The SUNRISE project introduces the SUNRISE Data Framework to link multiple SCDBs, requiring database owners to adhere to shared formats and abstraction levels to connect. This federated approach ensures unified access while maintaining flexibility for individual database management.

Together, these processes enable a structured, standardized, and scalable approach to scenario management for CCAM system safety assurance.

The following sections provide further details on each sub-block of the Scenario block.

4.1.1 Create

The Create block entails acquiring the necessary data and knowledge to create scenarios, while Format involves structuring the scenario using for example a scenario description language such as ASAM OpenSCENARIO XML [25] in combination with ASAM OpenDRIVE [26], or a schema based on the StreetWise domain model [12] or MetaScenario [61]. Finally, Store entails storing the formatted scenario in a searchable SCDB. This component also includes the SUNRISE Data Framework for accessing individual scenario databases.

Please note that the processes described in the Create and Format blocks, take place within each SCDBs. Within the Store block, the SUNRISE Data Framework is introduced to link individual SCDBs into the SUNRISE Data Framework, providing that these SCDBs meet the requirements identified and developed in the SUNRISE project.

Since the scenario creation takes place in individual SCDBs, the information provided here is for reference. However, from safety assurance point of view, the scenario creation process is an integral part. As mentioned earlier, scenarios can be created using two different approaches: data-based, and knowledge-based [17], however this does not restrict the total number of created scenarios as this is specific to the databases and also the use cases.

The Safety Pool™ scenario database [13], for example, contains scenarios generated using eight different methods which belong to the two approaches. StreetWise [12] uses driving data to extract real-world scenarios, determine the statistics of these scenarios, and use the scenarios for assessing automated driving systems. To support the research and development of StreetWise, it comes with a concept database that contains real-world scenarios. The EU ADS Act [22], FRAV VMAD [23], CertiCAV [24] also mention scenario creation approaches from both data-based and knowledge-based; SAKURA [11] mainly uses data-based scenario creation approach; StreetWise [12] also incorporate real world data combined with sampling methods to create scenarios; VVM [62] documents a data based scenario generation approach from drone database.

4.1.2 Format

Once the scenarios are created, the next step is to **Format** them in a way that represents the scenario effectively. Please note that since this activity takes place at individual SCDBs, there is no specific requirements on the format at this stage (as such requirement will be incorporated at later processes). The following information is for references only. Given the multiple stakeholders involved in scenario-based testing, including regulators, research engineers, test engineers, system engineers, and the public, there is a need to use multiple levels of abstraction for scenario formatting. There are four levels of scenario abstraction: *functional level scenarios, abstract level scenarios, logical level scenarios,* and *concrete level scenarios*. Each level emphasises different properties; for example, functional and abstract levels may focus on human readability, while logical and concrete levels may focus on machine readability. The importance of using common formats is important, such as ASAM OpenSCENARIO [25] and OpenDRIVE [26] formats for logical and concrete representations, and BSI Flex 1889 [37] for abstract-level representation.

4.1.3 Store

After scenarios are created and formatted, the next step is to **Store** them in a scenario database for downstream testing activities. The SUNRISE SAF proposes a common

SUNRISE Data Framework (DF) for accessing multiple scenario databases. Requirements have been set out for these databases to connect to the SUNRISE DF. Scenario creation and formatting is the responsibility of individual database owners. However, it is also their responsibility to ensure that these requirements are met to connect their databases to the SUNRISE DF, such as using a common scenario format and incorporating multiple abstraction levels.

4.2 Environment

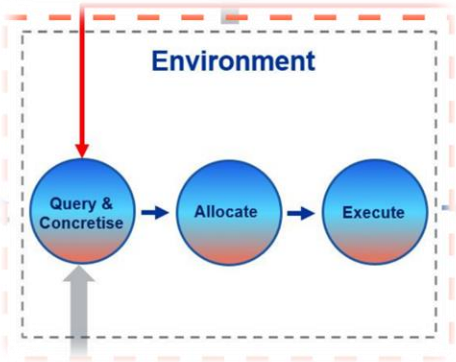

Figure 15 : Environment block of the SUNRISE SAF

The **Environment** block, which considers the Test Environment within the SUNRISE SAF operationalises test scenarios through three stages: querying and concretising, allocating test scenarios to test environments (virtual simulation, test track testing, or field operational testing), and executing the test scenarios, see Figure 15.

1. **Query & Concretise**
   Scenarios are retrieved from the SUNRISE DF using test objectives, ODD, behaviour, and external requirements. Logical scenarios, defined by parameter ranges, are concretised into specific values to create concrete test scenarios. Another option is to sample from the scenario categories that are available in the scenario database. The resulting test scenarios are then combined with test objectives to make them test-ready. Queries follow the OpenLABEL format [31], ensuring consistency across databases. Sampling methodologies support exploring parameter spaces, estimating safety measures, or identifying failure points.
2. **Allocate**
   Test scenarios are matched to appropriate test environments, from virtual simulations to controlled real-world settings. Test case requirements, such as scenario details and pass/fail criteria, are compared against the capabilities of available test instances like Hardware-in-the-Loop or proving grounds. A virtual-first approach prioritises lower-fidelity simulations initially, with iterative reallocations to higher-fidelity environments as needed. External factors, such as safety overrides, are documented for assessment.
3. **Execute**
   Allocated test scenarios are executed in the designated environments, with relevant data recorded for analysis. Feedback loops enable refining scenarios and optimising test objectives.

This component ensures a structured, efficient, and adaptive process for scenario testing within the safety assurance framework.

The following sections provide further details on each sub-block of the Environment block.

4.2.1 Query & Concretise

The **Query & Concretise** block takes input from the Input block, which contains a description of the ODD, the expected or required behaviour of the CCAM system, external requirements, and test objectives. It then passes these requirements to the SUNRISE DF as a query and retrieves scenarios from the individual scenario databases. The scenarios returned from the SUNRISE DF to the Query & Concretise block could be either of logical or concrete scenario levels. In case logical scenarios are returned, in a concretisation step, the parameters are defined within the applicable value ranges, allowing for a large number of concrete scenarios to be derived from a single logical scenario. These concrete scenarios with their specific parameter values are then combined with the test objectives and allocated within a test environment for execution.

Text below provides a workflow of the *Query & Concretize* block.

1. Since the scenario creation and formatting occurs in individual SCDB, the Input block information containing the test objectives, ODD & behaviour requirements and external requirements, will be fed directly into the Query & Concretise block.
2. These external requirements will then be used to query into individual SCDBs via the SUNRISE federated layer, and logical or concrete scenarios will be returned. It is at this step that scenarios hosted within SCDBs become test scenarios, because it is associated with the intended testing purposes upon the retrieval from the scenario databases.
3. If the returned scenarios are at the logical level (i.e., parameters are described in ranges)
    1. the first function of the block will create concrete scenario with concrete parameter values.
    2. the second function of the block will then combine the concrete test scenario with further test objectives.
4. If the returned scenarios are already concrete scenarios, then this block will skip the creation of the concrete scenario step, but combining it with test objectives.
5. The block will send the concrete test scenarios with their test objectives to the environment allocation block for test execution.
6. During or post execution, the Analyse block will then feedback the test outcome back to this block, the next set of concrete parameter combinations can be created. An example of such process could be incorporating an optimisation algorithm, with the intention of exploring a testing objective, e.g., to explore the failure points within a logical scenario.

Logical or concrete scenarios can be retrieved from databases connected to the SUNRISE Data Framework through queries. These queries are constructed using tags recorded in the OpenLABEL [31] format, which adheres to a harmonized ontology developed within the

SUNRISE project. This approach ensures a unified understanding of all elements and their interrelationships across the connected databases.

To derive concrete scenarios from the logical ones for testing purposes, several sampling methodologies have been developed. These methodologies facilitate the discretization of the continuous parameter space and enable the selection of specific samples (concrete scenarios) within this space. These samples are chosen to estimate the distribution of a safety measure across the parameter space. Alternatively, these methodologies can be applied to optimize for other testing objectives, such as identifying the pass/fail boundary within the parameter space or identifying parameter subspaces.

4.2.2 Allocate

The SUNRISE Safety Assurance Framework is test environment-agnostic, allowing scenarios to be executed in a range of test environments, from fully virtual to hybrid environments (such as Hardware-in-the-Loop) to controlled physical environments (such as proving grounds). Within the **Allocate** block, a test environment allocation workflow is applied (as detailed below), which will be implemented in case studies within the broader Safety Assurance Framework. Once a scenario is allocated to a test environment (also called test instance), the scenario is executed, and the corresponding data is recorded.

The initial allocation process involves two key inputs: test case information and test instance capabilities. Test cases include scenario descriptions, expected behaviour of the system under test (SUT), and pass/fail criteria. Specific requirements are extracted from these test cases. The second input consists of available test instances, such as virtual testing, X-in-the-Loop (XiL), and proving ground testing, with field testing excluded due to its uncontrollable nature.

The process begins with comparing test case requirements to the capabilities of test instances using a structured approach. This includes analysing aspects like scenery elements, environmental conditions, dynamic elements, and test criteria. Once a suitable test instance is identified, the test case is allocated to it.

A virtual simulation-first approach is prioritized for safety and efficiency. Test cases suitable for virtual simulation are executed using the lowest-fidelity simulation capable of meeting requirements, to maximize throughput. After execution, results are reviewed to determine if further testing on higher-fidelity test instances is required. This iterative process may include reallocations to ensure the necessary test coverage and accuracy.

The process includes provisions for external influence, such as road authorities overriding allocations or proving ground operators refusing tests due to safety concerns. These decisions must be documented for the final assessment. Additionally, the Allocate block includes initial reallocation steps, resulting from the Analyse block.

4.2.3 Execute

In the **Execute** block, the test execution might happen in a virtual, hybrid or physical test environment, depending on the test instance resulting from the Allocate block. The allocated test cases form the input for this block. How the tests are carried out is not explicitly specified by the SAF and is the responsibility of the entity performing the tests. If the Allocate block has been applied correctly, it is guaranteed that the selected test instance is capable of performing the tests. In the case of virtual testing, the SAF recommends a harmonised approach. For that, the SUNRISE project developed a harmonised V&V simulation framework, which can be used for virtual validation of CCAM systems but is not mandatory.

As shown in Figure 16, the **SUNRISE harmonised V&V simulation framework** consists of a so-called base layer consisting of 4 interconnected subsystems, namely:

- Subject Vehicle – Sensors (the sensors installed in the vehicle)
- Subject Vehicle – AD function (the behavioural competencies of the vehicle)
- Subject Vehicle – Vehicle Dynamics
- Environment (in which the vehicle operates)

In this approach, the base layer is the core element that can be harmonised, because these four subsystems are essential for all simulations. That is the reason why it is possible to use standardised interfaces between these subsystems. The framework can be extended by the user in four dedicated dimensions related to the target ODD, the vehicle Sensor set-up, the Software architecture and the Hardware architecture.

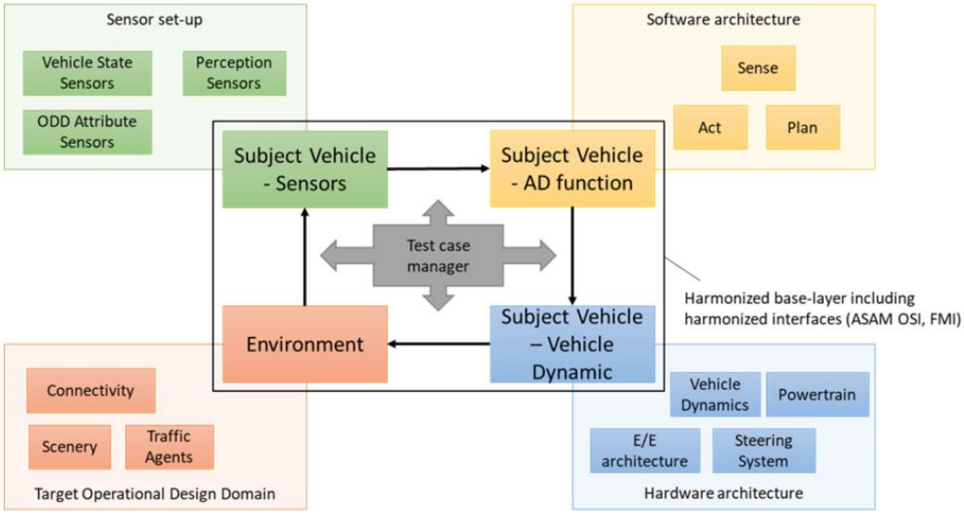

Figure 16: SUNRISE harmonised V&V simulation framework

4.3 Safety Argument

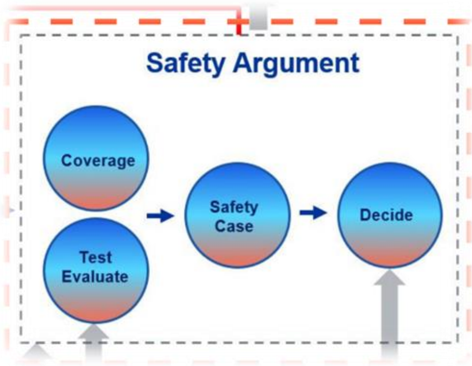

Figure 17: Safety Argument block of the SUNRISE SAF

The **Safety Argument** component evaluates test results to assess the system's safety through four interconnected stages: coverage analysis, test evaluation, safety case, and decision-making, see Figure 17.

1. **Coverage**
   This stage examines test coverage from multiple perspectives, including parameter ranges in logical scenarios and ODD features. Iterative feedback loops identify conditions where the system fails pass/fail criteria, enabling new concrete scenarios to be generated within the parameter space. ODD coverage is particularly critical, ensuring the system's operational boundaries are thoroughly tested using diverse scenarios.
2. **Test Evaluate**
   The results of individual test executions are assessed to determine whether the system meets safety requirements, such as maintaining speed limits or avoiding collisions. Results from both test evaluation and coverage analysis feed into iterative refinements of scenarios until coverage thresholds are achieved.
3. **Safety Case**
   The Safety Case block supports manufacturers in demonstrating that an ADS meets legal safety requirements by compiling structured evidence, including test results, safety arguments, and risk estimates. It addresses different types of regulatory requirements and emphasises the validation of methods, tools, and data used, forming a transparent foundation for regulatory assessment and deployment readiness.

4. **Decide**
   Once sufficient coverage is reached, combined results from the previous stages are synthesised to produce the overall safety assurance outcome for the system.

The Safety Argument component ensures systematic evaluation, iterative refinement, and comprehensive safety assurance of the tested system.

The following sections provide further details on each sub-block of the Safety Argument block.

4.3.1 Coverage

When referring to Coverage, it contains two aspect, for example 'X is covering Y', with this in mind bullets below provide a summary of the types of coverage considerations discussed in the SUNRISE project, followed by more detailed explanation.

- At the scenario set level: 1) ODD coverage by scenario set, 2) Requirements coverage by scenario set, 3) Behaviour coverage by scenario set
- At individual scenario level: logical parameter range covered by concrete parameter values

After execution, the data is analysed as part of the safety assessment of the system. There are two main blocks in this analysis: Coverage and Test Evaluate. The **Coverage** block examines the tests from multiple perspectives and derives a combined coverage outcome. These perspectives might include analysing coverage of the parameter value ranges in the logical scenario using individual concrete parameter values. Based on the analysis outcome, an iterative process derives new concrete scenarios within given logical scenario parameter ranges to identify failure conditions (e.g., parameter combinations that lead to system failure). This iterative feedback loop is indicated by the red feedback arrow.

Another aspect of Coverage Analysis is ODD analysis. Here, the accumulated ODD features covered by the set of test scenarios are examined to determine if enough of the system's ODD is covered. ODD is the focus because it defines the operational boundary within which the system is expected to operate safely. Thus, it is essential to thoroughly explore this claimed boundary using a diverse set of scenarios. These are just two examples that form the coverage concept, and the SUNRISE project will continue to explore this further in later stages.

4.3.2 Test Evaluate

The **Test Evaluate** block assesses each test execution to determine whether the system has passed or failed. For instance, did the system stay within an upper speed limit, or did it avoid collisions? Both the Coverage and Test Evaluate blocks contribute to the overall Analysis block, and the results are used to select further concrete parameters within the original scenario's parameter ranges. After several iterations, once the coverage threshold is reached, the combined coverage and test evaluation results will feed into the Decide block, producing the overall safety assurance outcome for the system. Please note that pass/fail criteria are heavily governed by the use case (i.e., use case dependent), therefore SUNRISE cannot prescribe specific criteria, however, the categories below should be considered: 1) whether the intended test has been executed (e.g., whether intended cut-in has occurred), 2) use case specific pass/fail criteria received from the input layer, 3) optionally, user can consider baseline based pass/fail criteria, such as human driver model baseline.

4.3.3 Safety Case

The SAF aims to support the development of a safety case for an Automated Driving System (ADS), enabling the manufacturer to determine whether the system is ready for assessment by a type-approval authority. When ready, the safety case is submitted to the authority, which decides whether there is sufficient evidence that the ADS meets the legal safety requirements and can be permitted on public roads.

Regulations and the Safety Case

Regulations often include different types of requirements (e.g. UNECE R157 [53] [63]). Each of these requirements and associated tests need to be included in the safety case:

1. Specific requirements – e.g. maintaining a minimum following distance (Clause 5.2.3.3.) – must be tested across a wide range of possible scenarios, typically with clearly defined metrics and pass/fail criteria. Test results demonstrating pass/fail in each test need to be provided, along with an explanation of scenario coverage relative to the ADS's ODD.

2. General requirements – such as the ADS shall not cause any reasonably foreseeable and preventable collisions (Clause 5.1.1.) – require not just quantified evidence, but reasoned argumentation supported by data showing what collisions can be considered reasonably foreseeable and that the ADS is able to avoid a collision in these scenarios.

3. Soft requirements – e.g. smooth and predictable vehicle behaviour (Clause 5.2.1.) – are less strictly defined and require qualitative or semi-quantitative assessments based on the demonstration of vehicle behaviour in a variety of scenarios. Documentation of system behaviour in relevant scenarios allows authorities to check such evidence, e.g. by evaluating the test reports and performing spot checks [64].

Quantifying Safety Risk

Despite the complexity of ADSs, safety assessment results must remain clear, transparent, and explainable to experts, politicians, and the public. A key metric is the residual safety risk associated with deploying an ADS on the road. Approaches such as Positive Risk Balance (PRB) and Globalement Au Moins Aussi Bon (GAMAB) are commonly referenced [65] [66]. Data-driven scenario identification allows for exposure levels and parameter distributions to be estimated. These scenarios feed into simulations that assess crash probabilities and consequences. The overall risk is calculated by combining crash probability, severity, and exposure across scenarios [67].

V&V of Methods, Tools, and Data

A credible safety case must be supported by evidence that the methods and tools used are valid and trustworthy. Documentation should cover:

- Data, scenario databases, and selection tools [68];
- Scenario generation and sampling methods;
- Equipment used in physical testing (e.g. vehicles, sensors, obstructions, measurement equipment);

- Simulation models (for vehicles, sensors, environment, etc.) [69].

If verification and validation (V&V) documentation is unavailable, appropriate validation efforts must be undertaken to fill these gaps. How to conduct V&V falls outside the SUNRISE project scope.

4.3.4 Decide

The **Decide** block serves as the culmination of the performance safety assurance, where all outputs generated through coverage and test evaluate are integrated into a clear and final decision: whether the system can be considered safe enough for approval and deployment. This decision is expressed as a **binary outcome**—pass or fail—regardless of the complexity and variability of preceding evaluation metrics. While earlier SAF blocks may involve graded performance indicators, confidence intervals, or probabilistic risk assessments, these must ultimately support a definitive, traceable, and auditable conclusion that aligns with regulatory expectations.

Residual risk—a key factor in the final pass/fail decision—arises from the gap between all possible real-world scenarios (the "yellow space") and those that have been tested (the "red space"), as seen in Figure 11. This gap reflects remaining uncertainty due to untested edge cases or underrepresented combinations of scenario parameters. The Coverage block quantifies this gap by statistical analysis or experts' domain knowledge. The Decide block then consolidates all available evidence to deliver a binary verdict—pass or fail. So to recap:

- Estimating the gap (how much) between scenario (yellow) and test scenarios (red) need to be assess in the Coverage
- Creation of the argument to justify the gap - Safety Case
- Acceptance of the argument is the task of the Decide

This approach is reflected in both regulatory and consumer-focused safety processes. In the case of ALKS (UN Regulation No. 157), the pass/fail decision is governed by explicit criteria—for instance, the vehicle must maintain a minimum time headway of 0.6 seconds and respond correctly to cut-in vehicles or emergency stops. Each test case has clearly defined thresholds, and failure to meet any criterion results in a binary fail outcome. On the other hand, Euro NCAP illustrates how a structured aggregation of many binary sub-decisions can feed into an overall performance score or safety rating. For example, a failure to detect and brake for a pedestrian in a night-time scenario is a binary fail within the AEB (Autonomous Emergency Braking) assessment, which contributes to the final star rating. In both cases, while the presentation of results may differ—approval vs. rating—the underlying logic is consistent: multiple test outcomes are assessed against pass/fail thresholds, and the accumulated results inform a final binary outcome on system.

5 SUNRISE DATA FRAMEWORK

The SUNRISE DF is a cloud application that serves as a federation layer to connect to external SCDBs. Although the SUNRISE Data Framework is not a mandatory part of the SUNRISE Safety Assurance Framework, its application is highly recommended since it provides the following benefits:

- It offers 1 single portal to access a wide variety of scenario databases through standardized interfaces
- It guarantees data governance through user authentication and authorization
- It enhances efficiency, accuracy, and compliance of scenario selection processes
- It centralizes scenarios for various stakeholders (like certifying entities, CCAM developers, test engineers)

The position of the SUNRISE Data Framework within the SUNRISE SAF, is shown in Figure 18 below. In this position, it can be observed as a tool to enhance overall CCAM validation efficiency through:

- Centralized scenarios from multiple databases through standardized interfaces
- Streamlined scenario selection enabled by advanced querying
- Improved compliance of selected scenarios with validation criteria
- Integration and compatibility with virtual, hybrid and physical test environments
- Sophisticated results tracking and feedback
- Future expansion of connected databases

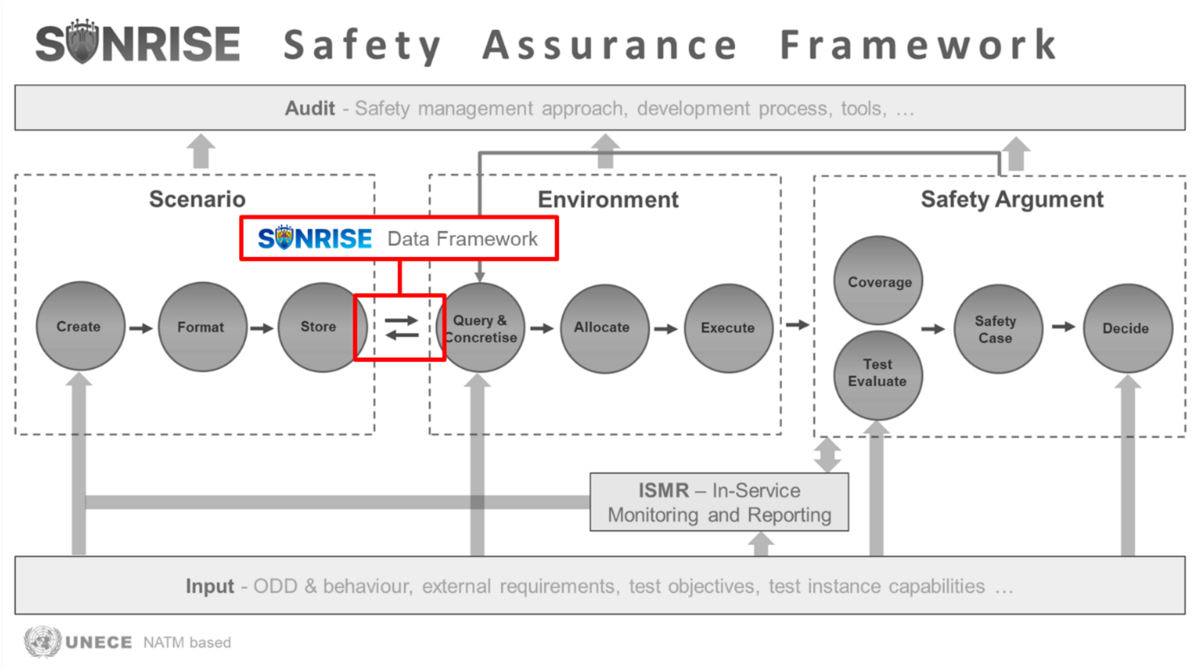

Figure 18: Position of the SUNRISE Data Framework within the SAF

Considering the complexity of the SUNRISE Data Framework, it is important to highlight the following principles to understand its scope:

- It functions as access point to external SCDBs rather than storing scenarios
- It requires SCDB host registration with API specifications
- It manages user credentials through automated processes
- It temporarily stores content during active user sessions
- It includes a Demo SCDB for illustration purposes

As detailed in Figure 19, the architecture of the SUNRISE Data Framework consists of a back-end framework that runs services and exposes APIs, interconnected with a front-end framework that serves web applications to users.

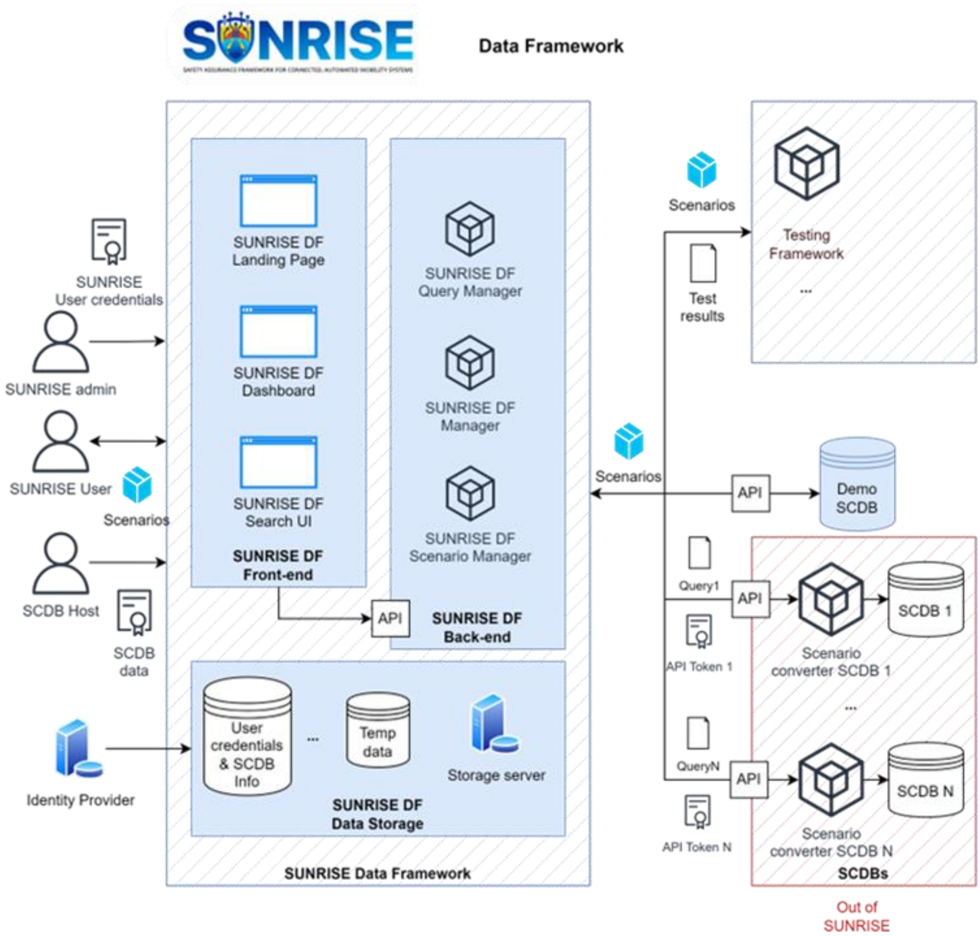

Figure 19: SUNRISE Data Framework architecture

The roles shown on the left side of Figure 19 above are defined as follows:

- **SUNRISE admin** manages the platform and its users, handling the activation of user accounts and validation of user onboarding requests. They are responsible for managing registered scenario databases and assigning user roles, which can be either SCDB Host or User level access.

- **SUNRISE user** accesses and utilizes SUNRISE Data Framework functionalities through the Dashboard and Search User Interface. All users must maintain valid credentials for specific scenario databases they wish to access.

- **SCDB host** is responsible for registering their scenario database in the SUNRISE Data Framework and providing the necessary API specifications and database details. They monitor their scenario database connection status and usage while ensuring database availability and proper integration with the framework.

- **Identity Provider** (IdP) manages user authentication and identity verification within the system. It provides secure tokens for user access and handles credential verification for scenario databases. The IdP serves as a critical security layer between users and scenario databases, enabling secure access to protected resources within the framework.

Further details of the front-end and back-end components of the SUNRISE Data Framework (as shown in Figure 19 above) are as follows:

1. **Front-end Components**

- SUNRISE DF Landing Page - This is where users select their preferred SCDBs to search for scenarios
- SUNRISE DF Dashboard - A front-end User Interface (UI) to visualize status of connected SCDBs and other statistics
- SUNRISE DF Search UI - Dedicated interface for triggering database searches
- SUNRISE DF Front-end - Integrated platform organizing front-end components

2. **Back-end Components**

- SUNRISE DF Query Manager - Manages queries (expansion, verification, formatting)
- SUNRISE DF Scenario Manager - Parses query results and validates formats
- SUNRISE DF SCDB connection - Handles user access management to SCDBs
- SUNRISE DF Back-end - Orchestrates components in cloud environment
- SUNRISE DF API - Exposes framework functions to its users, for example related to scenario search and management
- SUNRISE DF Data Storage - Manages internal databases for temporal and user data

The following **steps** indicate how a user can **apply** the SUNRISE **Data Framework**, and thereby enhance the overall CCAM validation efficiency of the SUNRISE SAF:

1. **Initial Access** (D6.1 Section 3.2.1):

   A. Login to SUNRISE Data Framework

   B. Access the SUNRISE Data Framework Dashboard (shown Figure 20 below)

   C. Select preferred SCDBs from the ones accessible to the user

2. **Scenario Search**:

   A. Use manual query through the graphical user interface (D6.1 Section 3.2.2)

B. Use automated query through the Automated Query Criteria Generation (AQCG) tool (D6.1 Section 3.2.3)

3. **Retrieve Scenarios** (D6.1 Section 3.2.4):

   A. Review returned scenarios from SCDBs

   B. Review metadata and relevant tags

4. **Validate Scenarios** (D6.1 Section 3.2.5):

   A. Use the Scenario Manager to validate retrieved scenarios

   B. Verify OpenSCENARIO and OpenDRIVE file formats

   C. Review error messages if any validation issues occur

5. **Test Environment Integration** (D6.1 Section 3.2.6, 3.2.7):

   A. Transfer validated scenario files to test environment

   B. Execute tests

   C. Retrieve test case results

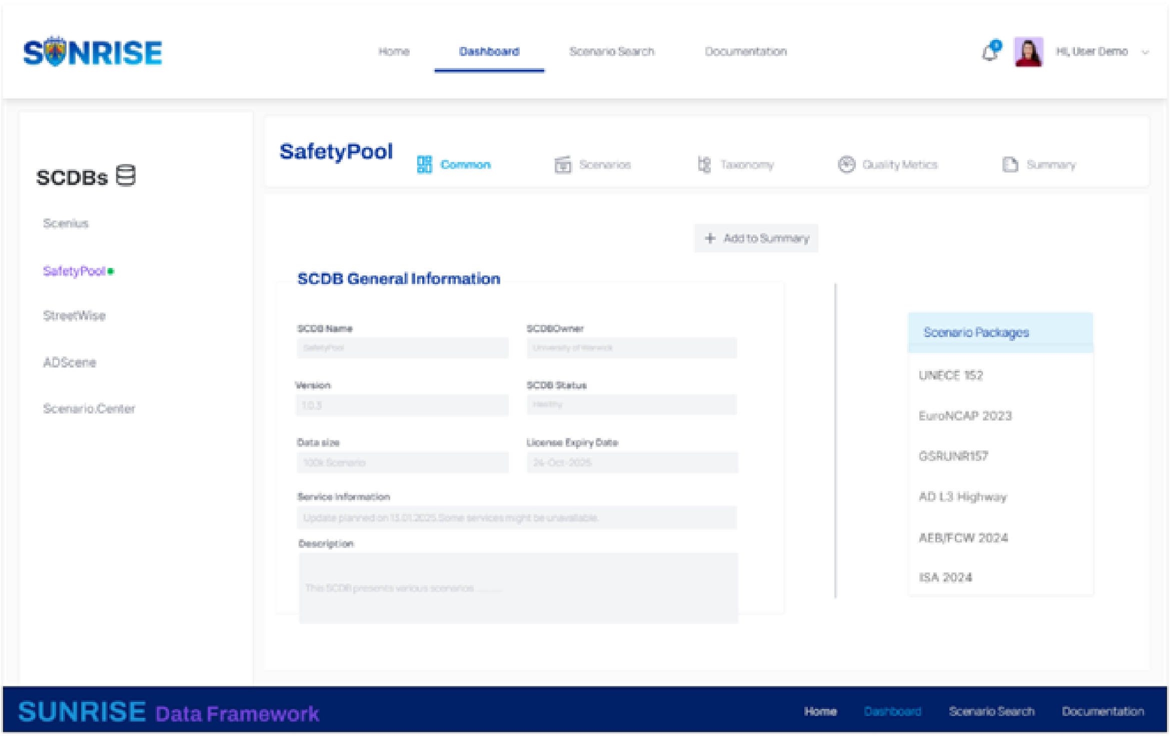

Figure 20: SUNRISE Data Framework Dashboard

6 HOW SPECIFIC TOPICS ARE ADDRESSED

This chapter outlines how the SUNRISE SAF addresses several specific and emerging topics that are critical to the development and deployment of CCAM systems. These include considerations such as SAE automation levels, human driver models, cybersecurity assessment, future adaptability, unknown scenarios, and applicability to CCAM functions based on artificial intelligence. The purpose of this chapter is to illustrate the flexibility and extensibility of the SAF by showing how these topics are accommodated within the framework's structure, thereby reinforcing its relevance across diverse technical, regulatory, and operational contexts.

6.1 SAE Levels

The SUNRISE SAF is a solution developed primarily for CCAM systems at SAE levels L3, L4 and L5. By focusing on these advanced automation levels, the SAF addresses the most complex safety challenges in vehicle automation, making it an essential tool for organizations working on the safety of CCAM systems (such as manufacturers, providers, certifiers and regulators).

While optimized for L3, L4 and L5 systems, the SAF can also be used to assess the safety of lower automation levels like SAE levels L1 or L2. In fact, the SAF has been validated in part by demonstrators operating at SAE level L2, providing practical insights into the SAF's application in real world state-of-the-art CCAM systems. The SAF's emphasis on higher automation levels ensures comprehensive assessment of safety-critical aspects where vehicles assume full driving responsibility (like in L4 and L5 systems). Although the SAF addresses safety related to driver-vehicle interactions (especially crucial for L3 systems), this matter is expected to be considerably improved in future updates of the SUNRISE SAF.

The SUNRISE SAF stands at the forefront with its robust, extendable and future-proof solution for CCAM safety assurance. Its comprehensive focus on higher automation levels (L3 and higher), combined with the flexibility to address lower levels (L2 and lower), makes it an invaluable solution for advancing the safe development and deployment of CCAM systems.

6.2 Human Driver Reference Models

Human driver reference models are an emerging element within the SAF. They aim to replicate or approximate human driving behaviour across different traffic and environmental contexts. Their integration into the SAF has been considered, particularly through collaboration with the i4Driving project [39].

Within the SAF, human driver models can serve three purposes: (1) to simulate surrounding traffic behaviour, (2) to provide a baseline for benchmarking the SUT, and (3) to represent the human element in SAE Level 2, 3 or 4 systems. The benchmarking use case is the most developed and was discussed during the SAF workshops. In this context, the model serves within the Test Evaluate block of the Safety Argument pillar, supporting comparisons between the SUT and a competent human driver operating under similar conditions.

Rather than developing in-house models, the SAF references established outputs from projects like i4Driving. That project offers validated behavioural models structured around core driving tasks and contextual factors. These models can act as a credible benchmark for regulators and developers during scenario-based testing. As the modelling landscape evolves, SUNRISE will continue to assess the role of human driver models in coordination with i4Driving and potentially other projects such as V4SAFETY [70]. In line with the Grant Agreement, internal development effort will be limited, with priority given to enabling external models to be applied within the SAF performance evaluation process. At the time of writing this report, the i4Driving model is still under development, further information can be found within the future deliverable (D4.5) of i4Driving work package 4.

6.3 Cybersecurity

Cyber vulnerabilities, such as spoofed sensor data, V2X message manipulation, or software tampering, can have direct safety implications and are thus relevant across multiple SAF components. Within the **Scenario** block, such threats can be modelled as part of scenario logic or environmental conditions. For example, a scenario may involve delayed or falsified infrastructure messages, simulating a degraded or malicious communications environment. These inputs enable targeted testing of the system's resilience to cyber-induced disturbances. In the **Execute** block, simulation and test environments can be configured to inject faults or disturbances reflecting cybersecurity failures, such as corrupted inputs or interrupted data flows. SAF's emphasis on execution traceability and tool credibility ensures that such tests maintain evidential value. In the **Safety Argument** block, cybersecurity test outcomes can be used to support claims about system robustness. Internal monitoring data, such as intrusion detection alerts or fallback activations, can serve as evidence of safe system responses under attack or failure conditions. By structuring cybersecurity concerns into scenario design, execution, and evaluation, SAF enables integrated testing of both functional and cyber-resilient safety performance. The SELFY [38] project concluded in May 2025, the information of the various tools which SELFY offers can be found at: https://selfy-project.eu/selfy-toolbox/, they cover the three top level areas: Cooperative Resilience and Healing System (CRHS), Situational Awareness and Collaborative Perception (SACP), Trust Data Management System (TDMS).

6.4 Future proofness

The SAF has been developed aiming to remain relevant beyond the project's end, taking into account foreseeable changes in technologies, regulations, and operational expectations. To support long-term relevance, several strategic actions are already in place. These include active involvement in international standardisation activities and policy discussions (e.g. UNECE, ISO, SAE, ASAM), and sustained alignment with international vehicle safety bodies. The SAF has also been designed to support wide adoption by its target users, such as OEMs, Tier 1s, and regulatory bodies, through stakeholder feedback loops.

Further future-proofing measures include the creation of exploitation plans, ensuring that key SAF components, such as scenario sets, metrics, KPIs, and the online handbook, can evolve in response to lessons learned from real-world deployment. In addition, future EU-funded projects such as SYNERGIES [71] and CERTAIN [72] are expected to build upon SUNRISE results, expanding the SAF's applicability. The breadth of the SUNRISE consortium and its participation in ongoing and future projects provides a strong foundation for maintaining the SAF as a living framework.

Another important item assisting in the uptake of the SAF beyond the end of the SUNRISE project, is the AI-based InfoBot integrated into the online SAF Handbook. This InfoBot can be consulted by SAF users up to 3 years after the project end, taking its answers from the combined pool of SUNRISE deliverables.

6.5 Unknown scenarios

Unknown scenarios pose a fundamental challenge to safety assurance, particularly in validating that a CCAM system behaves safely within its declared ODD. As such, the SAF incorporates several mechanisms aimed at reducing the likelihood of missing critical scenarios and improving the system's exposure to potentially unsafe, previously unseen conditions.

Within the SAF, the ISMR block plays a key role here. By requiring manufacturers to conduct significant public road driving and report safety relevant incidents and anomalies, the ISMR process helps surface unknown scenarios that may not have been captured through conventional scenario generation. This allows for iterative updates to both the SCDBs and internal safety assessment workflows. The intent is not simply to record these scenarios but to ensure that CCAM systems are sufficiently exposed to them and assessed for safe handling, a key expectation for Vehicle Safety Bodies.

The Safety Argument block supports this further through the Coverage block. It includes parameter range coverage within logical scenarios and ODD feature coverage across the scenario set. They quantify how much of the known parameter space is tested and can reveal gaps where unknown scenarios are more likely to arise. The SAF also enables random and synthetic scenario generation to explore untested combinations. In the Query & Concretise block, logical scenarios are used to generate new concrete scenarios through parameter sampling, while in the Create block, new scenarios can be synthesised if existing ones do not adequately cover the ODD. Although synthetic generation is part of the SAF's conceptual architecture, its detailed implementation lies outside the SUNRISE scope.

6.6 Role of human driver

The role of the human driver is a critical consideration in the safety assurance of CCAM systems. Automation changes rather than removing the human element, introduce new roles such as supervisor, fall-back user, remote operator, passenger, or other traffic participant. These roles bring variability and unpredictability that can significantly influence overall system safety, particularly in shared control situations or human-machine interaction. In the SAF, this topic is addressed in the **Execute** block. When the SUT includes human involvement, such as active driving or fallback intervention, the testing setup must reflect realistic human behaviours and their interaction with the system. This affects scenario design, system response assessment, and pass/fail criteria. The SAF remains system-agnostic but accommodates test designs where the human plays an integral role in the system's operation.

This topic is distinct from the use of human driver reference models for benchmarking system behaviour, which is addressed separately within the Test Evaluate block. Instead, the focus here is on how actual humans interact with the CCAM system under test, particularly in simulation and real-world execution environments. This includes aspects such as HMI, user response timing, trust, and decision-making. Relevant work has been carried out within SUNRISE under Task 4.6, which incorporates human actors in the simulation framework. These ensure that the SAF supports the representation and evaluation of human roles during test execution.

6.7 Artificial Intelligence

The SAF is designed to be technology agnostic, enabling safety assurance of CCAM systems regardless of the internal methods, architectures, or algorithms employed. This includes systems that incorporate AI, such as machine learning based perception modules, decision-making components, or end-to-end learning frameworks. Rather than tailoring assessment processes to specific technologies, the SAF adopts a black-box approach, treating the SUT as an opaque entity whose safety is judged solely on observable behaviour and performance across a representative scenario set.

This approach ensures that safety assessments remain consistent and impartial, even as CCAM systems evolve to include increasingly complex components. The SAF does not require visibility into internal algorithms or training data; instead, it relies on scenario-based testing, measurable safety performance, and pass/fail criteria derived from regulatory and operational requirements. This structure guarantees that AI-driven systems are assessed using the same processes as non-AI systems, maintaining fairness and robustness. The responsibility for selecting test scenarios that may expose potential weaknesses or edge behaviours in AI-based systems rests with SAF users. While the SAF provides the procedural framework and necessary tools (e.g. scenario allocation, test execution, analysis), tailoring the scenario set to reflect known or suspected AI limitations is outside the direct scope of SUNRISE.

7 SAF APPLICATION GUIDELINES

The following **steps** provide structured **guidance** for **applying** the SUNRISE **SAF** across its core functional blocks. Each step is supported by references to relevant SUNRISE deliverables (indicated by "D") for further detail and clarification. This stepwise format is intended to assist users in navigating the SAF implementation process in a methodical and traceable manner. While the structure may also support certification authorities or third-party assessors in evaluating SAF application, the emphasis here is on practical execution by those actively conducting the safety assurance process.

It should be noted that, while this guidance is focused on the operational elements of SAF application, it does not extend to internal audits of an organisation's broader development processes, such as lifecycle safety management, governance structures, or internal tool validation. These topics, though relevant to comprehensive safety oversight, fall outside the immediate scope of this guidance.

7.1 Scenario

The steps outlined below guide users in applying the **Scenario block** of the SUNRISE SAF, with a focus on evaluating the structure, quality, and completeness of a scenario database. These steps address the overarching characteristics of the scenario set rather than the detailed content of individual test cases, which are handled within the Environment block. It is important to note that the sub-processes related to scenario creation, formatting, and storage are typically the responsibility of the scenario database provider. As such, this guidance applies to the Scenario block as a whole and does not prescribe stepwise procedures for each sub-block individually.

1. **Check standardisation compliance:**
    - A. Verify ASAM OpenSCENARIO, OpenDRIVE, and OpenLABEL format compliance (D6.2 Section 3.3 and 1.5.2).
    - B. Confirm RESTful API compatibility for data exchange (D6.3 Section 3.2.3.5).
    - C. Check compliance with SUNRISE ontology standards (D5.2 Section 3.4.2; D6.2 Section 3.2).
2. **Evaluate quality of scenario set:**
    - A. Testing Purpose Metrics: Evaluate scenario relevance for specific test objectives (D5.3 Section 3.1, 4.1, 5.1, 6.1).
    - B. Scenario Description Metrics: Assess completeness and clarity of scenario descriptions (D5.3 Section 3.2, 4.2, 5.2, 6.2).
    - C. Exposure Metrics: Validate real-world representativeness and frequency data (D5.3 Section 3.3, 4.3, 5.3, 6.3).
    - D. (Dis)similarity Metrics: Ensure scenario diversity within the database (D5.3 Section 3.4, 4.4, 5.4, 6.4).
    - E. Coverage Metrics: Analyse parameter space coverage and ODD completeness (D5.3 Section 3.5, 4.5, 5.5, 6.5).

3. **Evaluate integration into SUNRISE Data Framework** (DF):
    A. Confirm proper connection to DF (D6.3 Section 2.3 and Annex 1).
    B. Verify authentication and authorisation mechanisms are properly configured for each scenario database (D6.2 Section 2.4, D6.3 Section 3.2.3).
    C. Verify if DF's query capabilities work as intended across connected scenario databases (D6.2 Section 2.2.2; D6.3 Section 2.3.1.3 and Annex 2).
    D. Validate semantic search functionality using harmonised ontologies (D6.2 Section 2.2.2; D5.2 Section 3.4).
    E. Confirm Automated Query Criteria Generation (AQCG) compatibility (D6.1 Section 3.2.3, D6.3 Section 3.2.5).
    F. Ensure scenario data exchange follows ASAM OpenSCENARIO, OpenDRIVE, and OpenLABEL standards (D6.2 Section 3.3).
    G. Validate input/output interface compliance with SUNRISE DF specifications (D6.2 Section 3.3).
    H. Confirm data format consistency across all connected databases (D6.2 Section 3.3).
4. **Verify source documentation:**
    A. Check that scenarios in the databases include source information (D3.2 Section 4.2.3).
    B. Validate that sources are properly documented with data- or expert knowledge origin (D3.2 Section 4.2.3, 8.1.1).
    C. Check exposure data availability for statistical analysis (D5.3 Section 4.3).
5. **Evaluate database extension capabilities:**
    A. Verify the database supports extensible parameter lists (D3.2 Section 5.5.2).
    B. Ensure the database can accommodate new parameters. For example when standards or protocols are updated. (D3.2 Section 5.5.2, 8.2.4).
    C. Verify that database supports parameter ranges (D5.3 Section 4.2).
6. **Verify database search and query functionality:**
    A. Confirm ODD-based filtering capabilities (D6.2 Section 2.2.1).
    B. Verify tag-based searching with standardised taxonomy (D3.2 Section 4.2.2).
    C. Validate omission-based queries (NOT statements) (D6.2 Section 1.5.2 and 2.2.1).
    D. Test query reproducibility and result traceability (D6.2 Section 1.5.2).
7. **Review data accuracy** (D5.3 Section 4.6):
    A. Assess the correctness of data entered into the scenario database.
    B. Cross-reference with known benchmarks or references.
    C. Verify that scenarios reflect real-world conditions accurately.
8. **Review data consistency** (D5.3 Section 4.6):
    A. Check for uniformity across all scenarios, for example in:
        - Units of measurement
        - Data formats
        - Terminologies used
        - Standardised naming conventions
9. **Check data freshness** (D5.3 Section 4.6):
    A. Check how up-to-date the scenarios are.

   B. Verify current relevance of the database content.
   C. Check last update timestamps.
10. **Check number of scenarios** (D5.3 Section 4.6):
    A. Count total distinct scenarios available.
    B. Obtain overview of database comprehensiveness.
    C. Check scenario quantity metrics.
11. **Check covered kilometres** (D5.3 Section 4.6):
    A. Check on how many KM's the scenario database is based.
    B. Assess span and scale of included scenarios.
    C. Quantify geographic coverage.
12. **Verify scenario distribution** (D5.3 Section 4.6):
    A. Analyse breakdown of scenarios by various categories such as:
       - Geographic regions
       - Road types
       - Weather conditions
       - Time of day
       - Traffic density
13. **Verify scenario complexity** (D5.3 Section 4.6):
    A. Assess difficulty levels across scenarios considering:
       - Number of vehicles involved
       - Presence of pedestrians
       - Road complexity
       - Environmental conditions

7.2 Environment

7.2.1 Query & Concretise

By following the steps outlined below, users of the SUNRISE SAF can apply the **Query & Concretise block** to ensure that the resulting test cases are of high quality, relevant to the system under test, and comprehensive enough to support the intended safety argumentation.

1. **Review the inputs** to the Query & Concretise block, previous called COTSATO process (COncretsing Test Scenarios and Associating Test Objectives) (D3.2 Section 7.3):
   A. Verify that the ODD description is provided and follows the format guidelines in ISO 34503.
   B. Check that the system requirements are clearly defined.
   C. Ensure the CCAM system under test is properly specified.
   D. Confirm that variables to be measured during test execution are listed.
   E. Validate that pass/fail criteria for successful test execution are defined.
2. **Evaluate use of SUNRISE Data Framework** (DF):
   A. Confirm proper connection to federated scenario databases through DF (D6.3 Section 2.3 and Annex 1).
   B. Verify authentication and authorisation mechanisms are properly configured for each scenario database (D6.2 Section 2.4, D6.3 Section 3.2.3).

C. Verify if DF's query capabilities work as intended across multiple connected scenario databases (D6.2 Section 2.2.2 and 3.1; D6.3 Section 2.3.1.3 and Annex 2).
D. Validate semantic search functionality using harmonised ontologies (D6.2 Section 2.2.2; D5.2 Section 3.4).
E. Confirm Automated Query Criteria Generation (AQCG) compatibility (D6.1 Section 3.2.3, D6.3 Section 3.2.5).
F. Ensure scenario data exchange follows ASAM OpenSCENARIO, OpenDRIVE, and OpenLABEL standards (D6.2 Section 3.3).
G. Validate input/output interface compliance with SUNRISE DF specifications (D6.2 Section 3.3).
H. Confirm data format consistency across all connected scenario databases (D6.2 Section 3.3).

3. **Examine the query** used to fetch scenarios from the scenario database(s):
   A. Verify Ontological Compliance (D5.2 Section 4 and 3.4.3):
      - Ensure query uses the SUNRISE ontology structure.
      - Confirm query operators are properly implemented (AND, OR, NOT, ..).

   B. Apply Standardised Data Format Requirements (D5.2 Section 4):
      - Verify query compatibility with OpenLABEL JSON format for scenario tagging.
      - Ensure query supports OpenSCENARIO and OpenDRIVE format requirements.
      - Validate that query can handle both logical and concrete scenario representations.
4. **Evaluate quality of individual scenarios:**
   A. Check the testing purpose metrics to ensure scenarios are relevant for the intended testing (D5.3 Section 3.1, 4.1, 5.1, 6.1).
   B. Verify scenario description quality including completeness and unambiguity (D5.3 Section 3.2, 4.2, 5.2, 6.2).
   C. Assess scenario exposure and probability to verify if scenarios represent realistic situations (D5.3 Section 3.3, 4.3, 5.3, 6.3).
   D. Assess diversity and similarity between scenarios to avoid redundant testing (D5.3 Section 3.4, 4.4, 5.4, 6.4).
   E. Verify scenario coverage to ensure comprehensive testing of the Operational Design Domain (ODD) and the parameter space (D5.3 Section 3.5, 4.5, 5.5, 6.5).
   F. Check that the scenario concept complies with the requirements (D3.2 Section 4).
   G. Confirm that the scenario parameters meet the requirements (D3.2 Section 5).
   H. Validate that the parameter spaces adhere to the requirements (D3.2 Section 6).
5. **Review application of the subspace creation methodology** required to convert logical scenarios (obtained through querying) into concrete test cases for execution:
   A. Identification of critical parameter regions (D3.4 Section 4.1 and 4.6).
   B. Structured exploration of the parameter space (D3.4 Section 4.2).

C. Focus on performance-sensitive regions (D3.4 Section 4.6).
D. Identification of multiple failure conditions (D3.4 Section 4.7).
E. Focus on pass/fail boundaries rather than exhaustive parameter coverage (D3.4 Section 4.8).

6. **Review the test cases** generated by the COTSATO process (D3.2 Section 7.5):
    A. Ensure each test case includes a test scenario, metrics, validation criteria, and pass/fail criteria.
    B. Depending on the purpose, verify that the metrics cover aspects such as safety, functional performance, HMI, operational performance, reliability, and scenario validation.
    C. Confirm that a clear mapping exists between system requirements and the generated test cases.

7.2.2 Allocate

By following the steps outlined below, users of the SUNRISE SAF can apply the **Allocate block** to ensure that the allocation of test cases to appropriate test environments is performed correctly and comprehensively, in accordance with the SAF's recommended procedures and decision logic.

1. **Review the comparison** of test case requirements with test instance capabilities:
    A. Ensure that the structure outlined in D3.3 Section 3 was followed, which includes scenery elements, environment conditions, dynamic elements, and test criteria (D3.3 Section 3.3).
    B. Verify the consideration of various metrics:
        - Check that both functional and non-functional metrics were considered (D3.3 Sections 4.3 and 4.4).
        - Confirm that safety was prioritised in the decision-making process (D3.3 Section 4.5).
    C. Verify that specific requirements of the system under test (SUT) are considered alongside test case requirements (D3.5 Section 4.1.3).
    D. Check for iterative refinement of both test case requirements and test instance capabilities based on initial allocation outcomes (D3.5 Section 4.2).
2. **Review the initial allocation process**:
    A. Confirm that the process outlined in D3.3 Section 4.5 and D3.3 Figure 27 was followed.
    B. Verify that safety standards such as SOTIF were considered in the allocation process, particularly for identifying potentially triggering conditions or functional insufficiencies of the SUT (D3.3 Section 4.5).
    C. Validate scenario realisation by (D3.5 Section 5.1):
        - Verify that allocated scenarios can be meaningfully realised at the semantic level, not just technically executed.
        - Confirm that the SUT will encounter intended triggering conditions and respond appropriately according to scenario logic.
        - Ensure all relevant behavioural phases can be reached in the allocated test environment.

D. Verify if special circumstances resulted in deviations from the general methodology. And if so, ensure they were properly justified (D3.3 Section 4.5).

3. **Review the re-allocation process** (when necessary):
    - Ensure that the iterative re-allocation to higher-fidelity test instances, was performed as described in D3.3 Section 4.5.
    - Verify that the reasons for re-allocation decisions were properly documented (D3.3 Section 4.6).
    - Check if safety confidence evaluation has been performed for scenarios where uncertainty exceeds acceptable false acceptance risk (D3.5 Section 5.2.1).
    - Check if selection strategies for higher-fidelity validation have been applied (D3.5 Section 5.2.2). For example through:
        - Integration of higher-accuracy vehicle dynamics models
        - Adoption of physics-based sensor models
        - Implementation of co-simulation frameworks
        - Model improvement based on correlation against real-world data
    - Check if correlation analysis has been performed to validate alignment between different fidelity levels (D3.5 Section 5.2.3).
4. **Verify expert evaluation** in case the correlation across SAF workflow components reveals discrepancies (D3.5 Section 5.3). For example, check for issues related to:
    - Input block - Insufficient ODD or behaviour descriptions.
    - Scenario block - Mismatched scenarios and queries.
    - Environment block - Allocation or execution problems.
    - Technical simulation issues - Sensor models, vehicle dynamics, environment simulation.
    - Documentation of correlation root causes (for future allocation improvements) and updates to test instance capability knowledge.
5. **Examine the documentation** of the (re-)allocation results:
    - Check for the presence of a tree structure containing all metrics and results of the comparison to all test instances (D3.3 Section 4.6).
    - Ensure that all steps of the (re-)allocation process, including reasons for decisions and selection criteria used, were documented and returned to the SAF (D3.3 Section 4.6, 5 and Figure 28).
    - Ensure that scenarios that could not be (re-)allocated or were not sufficiently tested are properly flagged and reported to the Coverage block of the Safety Argument component of the SAF (D3.3 Section 4.6 and Figure 28).
    - Check for documentation of all safety confidence evaluations and their outcomes (D3.5 Section 5.2.1).
    - Check for documentation of correlation results between different fidelity levels (D3.5 Section 5.2.3).
    - Check for documentation of scenario realization validation results (D3.5 Section 5.1).
    - Check for documentation of iterative refinement of test case requirements and test instance capabilities (D3.5 Section 4.2).

6. **Review the integration with the SAF workflow:**
   A. Check for existence of feedback mechanisms to receive test execution outcomes from subsequent SAF blocks (D3.5 Section 3 and 5.2).
   B. Check for existence of refinements of test case requirements and test instance capabilities, resulting from executed test case results (D3.5 Section 4.2).
   C. Ensure that validated test outcomes contribute to trustworthy safety arguments (D3.5 Section 6).

7.2.3 Execute

By following the steps outlined below, users of the SUNRISE SAF can apply the **Execute block** to ensure that the simulation framework used, meets the essential requirements and delivers reliable, validated results. While the use of the SUNRISE harmonised V&V simulation framework is recommended — particularly for its benefits in interoperability, modularity, and scalability — it is not a mandatory requirement for applying the SAF.

1. **Verify selection and validation of appropriate testing environments**:
   A. Check the assessment of virtual testing limitations (D4.6 Section 2.3), including:
      - Latencies in data exchange between co-simulation models.
      - Semantic gaps between real world and simulation (sim2real gap).
      - Sensor simulation discrepancies with real sensor behaviour.
   B. Check the selection of appropriate hybrid or real-world testing methods when serious virtual testing limitations are identified (D4.6 Section 2.4):
      - Black Box testing on proving grounds or public roads (D4.6 Section 3.1).
      - Vehicle-in-the-Loop testing for AD function validation (D4.6 Section 4).
      - Driver-in-the-Loop testing for human-machine interaction scenarios (D4.6 Section 5).
2. **Verify that the simulation framework used, aligns with the Harmonized V&V Simulation Framework** (D4.4 Section 4.5), which includes**:**
   A. Checking that the base layer contains the four (or less if applicable) core interconnected subsystems:
      - Subject Vehicle - Sensors
      - Subject Vehicle - AD function
      - Subject Vehicle - Vehicle Dynamics
      - Environment
   B. Validate the data formats used, align with recommended standards (D4.4 Section 4.3):
      - ASAM OpenSCENARIO for scenario descriptions
      - ASAM OpenDRIVE for road networks
      - ASAM OpenLABEL for sensor data and scenario tagging
   C. Confirming the framework uses standardised interfaces between subsystems, particularly ASAM OSI as detailed in D4.4 Section 4.4.
3. **Assess virtual test validation**:
   A. Check that appropriate validation setup has been applied (D4.5 Section 3):

    - A.1-A.3 for separate subsystem validation
    - B.1-B.2 for integrated validation
    - C.1-C.3 for comprehensive system validation
   B. Verify that correlation analysis between virtual simulation and physical tests was performed (D4.1 Section 3.6, D4.2 Section 8.1 - R1.1_14, D4.5 Section 3).
   C. Confirm robustness and representativeness of virtual validation framework (D4.2 Section 8.1 - R1.2_10).
   D. Check that model quality metrics meet defined thresholds (D4.1 Section 3.6, D4.5 Section 3.1).
   E. Review the simulation model validation test report (D4.1 Section 3.6).
4. **Assess hybrid test validation**:
   A. Assess the validation of virtual components (D4.6 section 4.2 and 6):
    - Verify all virtual models were validated before use in hybrid testing.
    - Check virtual environment validation for realistic representation.
    - Confirm calibration processes align virtual inputs with real sensor outputs.
    - Check for accounting of noise, latency, and distortion in sensor stimulation.
   B. For Vehicle-in-the-Loop tests on Proving Grounds (D4.6 Section 4.1):
    - Verify real vehicle with real AD function was used.
    - Check that virtual traffic agents and sensor models were properly integrated.
    - Confirm ETSI-compliant V2X messaging for collective perception testing.
   C. For Vehicle-in-the-Loop tests on Testbenches (D4.6 Section 4.2):
    - Check sensor stimulator calibrations (GNSS, Radar, Camera, LiDAR).
    - Verify correlation results.
    - Check real-time synchronisation between physical and virtual components.
   D. For Driver-in-the-Loop on Proving Grounds (D4.6 Section 5.2):
    - Verify VR/AR system integration with real vehicle.
    - Check optical tracking system with stereo camera pairs.
   E. For Driver-in-the-Loop in Simulators (D4.6 Section 5.3):
    - Verify HMI device integration.
5. **Assess physical test validation**:
   A. Assess Black Box Testing on Proving Grounds or Public Roads (D4.6 Section 3.1):
    - Verify system under test (SUT) proper instrumentation of cameras, sensors, and measurement systems.
    - Check implementation of two-step validation: 1st test tracks, 2nd public roads.
6. **Evaluate validation metrics and Key Performance Indicators (KPIs):**
   A. Verify that requirements from protocols, standards and regulations were used where applicable. For example, from Euro NCAP, GSR or 1958 agreement

(D4.2 Section 8.1 requirement R1.1_01, R1.2_01 and Section 8.1 requirement R3.1_01).

B. Check that SOTIF requirements were addressed (D4.2 Section 4.1), including:
   - Risk quantification for scenarios, triggering conditions and ODD boundaries.
   - Validation results for known unsafe scenarios.
   - Validation results for discovered unknown unsafe scenarios.
   - Assessment of residual risk.

7. **Check test case execution results:**
   A. Check that all executed test cases generated desired results.
   B. Confirm test coverage metrics have been generated. For example:
      - Check EURO NCAP and GSR compliance metrics (D4.2 Section 8.1 - R1.1_01, R1.2_01, R3.1_01).
      - Verify sensor validation metrics were applied (D4.2 Section 8.1 - R1.1_02).
      - Review correlation coefficients between simulation and physical test results (D4.2 Section 8.2 - R2.1_49).
   C. Confirm that test results include both virtual and physical validation data where applicable (D4.2 Section 8.1 - R1.1_14).
   D. Verify that executed simulations correspond to the requests from scenario manager (D4.2 Section 8.1 - R1.1_25).
8. **Verify documentation** related to test execution:
   A. Check for documentation of test results including key outcomes and values of metrics and KPIs, and ensuring that validated test outcomes contribute to trustworthy safety arguments (D3.5 Section 6).
   B. Check for documentation of the applied validation methodology and metrics, including the decision-making process (D4.5 Section 3 Introduction and 3.1).
   C. Check for documentation of values of correlation quality metrics (D4.5 Section 3.1).
   D. Check for documentation of toolchain functionality assessment results for repeatability, uncertainty, performance, and usability (D4.5 Sections 3.7-3.9).
   E. Check for documentation on matching use case requirements to validation setup (D4.5 Section 4).

7.3 Safety Argument

7.3.1 Coverage

By following the steps outlined below, users of the SUNRISE SAF can apply the **Coverage block** to ensure that the combined set of test case results provides sufficient coverage of the ODD and the relevant parameter value ranges necessary for a robust safety argumentation.

1. **Review ODD coverage** (D5.3 section 5.5.1):
   A. Assess Tag Coverage (D5.3 section 5.5.1.1):
      - Review how scenarios are categorised and tagged
      - Check that all important driving situations are represented

        - Verify each tag appears across different scenario types
        - Look for missing or underrepresented driving conditions
    B. Assess Time Coverage (D5.3 section 5.5.1.2):
        - Check that scenarios cover different time periods adequately
        - Identify any time gaps where no scenarios exist
        - Review scenarios that happen at the same time
        - Document any uncovered time periods
    C. Assess Actor Coverage (D5.3 section 5.5.1.3):
        - Review which vehicles and road users are included in scenarios
        - Check coverage at different distances from the test vehicle
        - Verify that vehicles in critical positions are well represented
        - Assess if important actors are missing
    D. Assess Actor-Time Coverage (D5.3 section 5.5.1.4):
        - Check that important actors appear throughout relevant time periods
        - Verify actors stay covered during their entire critical presence
2. **Review parameter space coverage** (D5.3 section 5.5.2):
    A. Review how selected parameter values are analysed for completeness
    B. Check that statistical assumptions are properly verified
    C. Identify areas with insufficient parameter coverage
    D. Verify that synthetic scenarios fill identified gaps
3. **Check SAF integration** (D5.3 section 3 introduction):
    A. Verify that coverage analysis supports scenario selection
    B. Check that coverage results inform the safety argumentation
    C. Ensure that coverage metrics support completeness arguments
4. **Final verification** (D5.3 sections 5.5 and 6.5):
    A. Confirm all coverage metrics meet acceptable levels
    B. Validate that coverage gaps are properly addressed
    C. Ensure synthetic scenarios appropriately fill coverage gaps
    D. Cross-check results with implementation examples

7.3.2 Test Evaluate

By following the steps outlined below, users of the SUNRISE SAF can apply the **Test Evaluate block** to ensure that trustworthy evidence is produced that can be reliably used in building the overall safety argumentation of the CCAM system under test.

1. **Verify test run validation** including proper application of *test run validation metrics*, to ensure that (D3.5 Section 5.2):
    A. The test execution was valid and meaningful
    B. The correct test instance was used for the scenario
    C. Test scenario importance was properly evaluated
    - Critical scenarios were appropriately prioritised
2. **Assess scenario realisation** to verify that the scenario was meaningfully executed, by checking that (D3.5 Section 5.1):
    A. The CCAM system actually encountered the intended triggering conditions

B. The CCAM system responded appropriately to the scenario's defined logic
C. The scenario behaviour phases that test the intended function, were reached
  - Any scenarios not properly realised were flagged as "not achieved"

3. **Evaluate safety confidence metrics** by verifying that (D3.5 Section 5.2.1):
   A. An acceptable false acceptance risk was properly defined
   B. Uncertainty estimation at the test point was established
   C. Uncertainty was compared against false acceptance risk thresholds
   D. Sufficient confidence margins were maintained
4. In case higher fidelity testing was performed, **review the correlation analysis** by checking that (D3.5 Section 5.2.3):
   A. Correlation between low and high fidelity test instances was examined
   B. Appropriate correlation methods were used (like the Pearson correlation coefficient)
   C. Statistical significance was verified (low p-value)
   D. Acceptable correlation levels were predefined before testing
5. **Validate expert analysis documentation** by ensuring it includes (D3.5 Section 5.3):
   A. Root cause analysis for correlation discrepancies
   B. Assessment of re-allocation needs to different test instances
   C. Documentation of further investigations when required
   D. Expert explanations for anomalies or unexpected results

7.3.3 Safety Case

By following the steps outlined below, users of the SUNRISE SAF can apply the **Safety Case block** to ensure that an Automated Driving System (ADS) has sufficient evidence-based documentation to demonstrate compliance with (legal) safety requirements and support regulatory type-approval decisions for deployment on public roads.

1. **Review regulatory requirements** which can be subdivided as follows (D2.3 Section 4.3.3):
   A. Specific Requirements:
      - Verify that test results demonstrate pass/fail in each test
      - Check for explanation of scenario coverage relative to the ADS's ODD

   B. General Requirements:
      - Review reasoned argumentation supported by data
      - Confirm demonstration of what collisions can be considered reasonably foreseeable
      - Verify that the ADS is able to avoid collisions in these scenarios

   C. Soft Requirements:
      - Assess qualitative or semi-quantitative assessments
      - Review documentation of system behaviour in relevant scenarios
      - Enable authorities to check evidence through evaluation of test reports and performance of spot checks
2. **Assess safety risk quantification** (D2.3 Section 4.3.3):
   A. Ensure assessment results are clear, transparent and explainable to experts, politicians, and public.

B. Review documentation of residual safety risk associated with ADS deployment.
C. Verify approaches used (such as Positive Risk Balance or GAMAB).
D. Confirm that data-driven scenario identification includes exposure levels and parameter distributions.
E. Validate that overall risk calculation properly combines crash probability, severity, and exposure across scenarios.

3. **Verification & Validation (V&V) of methods, tools, and data** by checking that documentation covers (D2.3 Section 4.3.3):
   A. Data, scenario databases, and selection tools
   B. Scenario generation and sampling methods
   C. Equipment used in physical testing (vehicles, sensors, obstructions, measurement equipment)
   D. Simulation models (for vehicles, sensors, environment, etc.)

   Important Note: If *V&V documentation* is unavailable, appropriate validation efforts must be undertaken to fill these gaps.

7.3.4 Decide

By following the steps outlined below, users of the SUNRISE SAF can apply the **Decide block** to ensure that test outcomes are assessed against pass/fail thresholds, and the accumulated results inform a final binary decision on the safety of the CCAM system under test, which is well-grounded, transparent, and defensible.

1. **Verify binary decision output** (D2.3 Section 4.3.4):
   A. Confirm that a clear pass/fail decision has been made about the CCAM system's safety for approval and deployment.
   B. Ensure the conclusion is definitive, traceable, and auditable.
2. **Audit residual risk assessment** (D2.3 Section 4.3.4):
   A. Check that residual risk has been properly identified.
   B. Verify the gap between all possible real-world scenarios (the "yellow space" in D2.3 Figure 11) and tested scenarios (the "red space" in D2.3 Figure 11) has been considered in the decision.
3. **Confirm decision logic integration** (D2.3 Section 4.3.4):
   A. Ensure all outputs from the Coverage and Test Evaluate blocks have been properly integrated into the decision.
   B. Verify that any graded performance indicators, confidence intervals, or probabilistic risk assessments ultimately support the binary decision.
4. **Review acceptance of safety case arguments** (D2.3 Section 4.3.4):
   A. Confirm that the Safety Case arguments justifying the gap between scenarios and test scenarios, have been properly evaluated.
   B. Verify that the acceptance of the Safety Case arguments has been properly documented.

8 SAF LIMITATIONS

As can be seen in previous chapters of this document, the SUNRISE SAF provides a solid basis for comprehensive safety assurance of CCAM systems. Nevertheless, users should be aware of the limitations of the SUNRISE SAF, since those might lead to incorrect decisions on the safety of CCAM systems, if not taken into account. For that reason, this chapter describes **all known limitations** of the SUNRISE SAF:

1. **Technology Readiness Level**

   Under SUNRISE, the technology readiness level (TRL) of the SAF has been raised to 6 or 7. It demonstrates a working prototype by means of various CCAM Use Cases. This demonstration also serves as an initial validation of the SAF. However, a full-scale validation of the SAF has not been done, due to scope limitations. SAF users should be aware of this limitation when applying the SAF. In practice this could mean:

   A. Having to divert from what the SAF proposes.
   B. Having to approximate what the SAF proposes.
   C. Having to apply alternatives to what the SAF proposes.
   D. Having to solve possible limitation(s) of the SAF.

2. **Scope limitations**

   Although integral parts of the SAF, due to limitations in the SUNRISE scope, certain SAF blocks are not fully developed. These blocks are:

   A. **Scenario** involves managing scenarios that are critical for safety assurance, consisting of three sub-blocks: Create, Format and Store. Together, these blocks enable a structured, standardized, and scalable approach to scenario management. Since these blocks are considered the responsibility of the scenario database owners, they are not further elaborated in the SUNRISE project.

   B. **ISMR** (in-service monitoring and reporting) captures and records additional information during system deployment, which can be considered for future system designs. It occurs during system operation, with continuous monitoring by the manufacturer or fleet operator for purposes of continual safety assessment and improvement.

   C. **Audit** focuses on ensuring that the development process, tools used, and overall safety management approach are adequate. It ensures that the manufacturer has proper processes in place for operational and functional safety throughout the system's development and lifecycle, confirming the safety of the system's design and sufficient validation before market introduction.

3. **Standardization challenges**

The SAF relies on multiple standards that are still in development or have been very recently published, creating challenges in synchronizing and establishing them as common practice. Furthermore, there is a lack of standardized metrics for CCAM performance and safety evaluation, including new metrics for emerging technologies like AI-based systems. International agreement is needed on ODD definition, scenario representation and safety metrics, along with the creation of global open-access scenario databases. This would enhance interoperability and consistency across different CCAM safety assurance approaches and stakeholders. Under SUNRISE important steps were taken towards that goal, but reconciling international standards and regulations involves significant efforts and difficulties, that remain a continuous challenge.

4. **Human driver reference models**

The SAF uses human driver reference models for comparison of safety performance between human drivers and CCAM systems. These models are not developed within SUNRISE but sourced from an external project (i4Driving). This creates a dependency and therefore involves certain risks that could negatively affect a successful comparison between human drivers and a CCAM system.

5. **Human-Vehicle interaction**

The SAF's emphasis on higher automation levels ensures comprehensive assessment of safety-critical aspects where vehicles assume full driving responsibility. The SAF also addresses safety related to human-vehicle interactions (especially crucial for L3 systems). However, this matter is expected to be considerably improved in future updates of the SAF, particularly regarding HMI interaction. For example, in the CERTAIN [72] project resulting from HORIZON-CL5-2024-D6-01-02.

6. **Cybersecurity**

Although one of its Use Case (UC4.2) partially covered cybersecurity aspects focusing on the 'Execute' part of the SAF, for safety assurance related to cybersecurity, the SAF refers to an approach developed in another European funded project (SELFY [38]) cybersecurity.

7. **Incident procedures**

To further improve the safety of CCAM systems, detailed procedures for incident response and recovery may be necessary (during testing and in-service monitoring). This includes how data from incidents is collected, analyzed, reported and used to improve system safety, as well as how the system should respond automatically in the event of malfunctions or accidents to minimize harm. However, this topic belongs to the ISMR block of the SAF, and is therefore not covered by the scope of the SUNRISE project.

8. **Numeric threshold values**

   Although guidelines are provided for the establishment of safety metrics (like pass/fail criteria), the SAF does not provide any numeric threshold values. The establishment of such values can vary from region to region, and is subject to cultural, societal and political influence. It is therefore not considered within the scope of the SUNRISE project to provide these values.

9. **Tool dependency**

   While the SAF is designed to be *tool-agnostic*, implementing it effectively requires users to select compatible tools that fit their specific needs.

9 CONCLUSIONS

The **SUNRISE SAF** presents a harmonised, modular, and future-proof approach to assess the **safety of CCAM systems**. Developed in alignment with the UNECE NATM [1] pillars and building upon scenario-based principles, the SAF integrates a structured and interoperable methodology and a federated data framework (the SUNRISE Data Framework).

During the initial design phase, thorough analysis was performed of existing safety assurance frameworks and a detailed survey of current international initiatives (section 2.1.1). Throughout its development, the SAF has maintained **international alignment** with activities from key standardisation organisations such as SAE, ASAM, UNECE, ISO, and BSI (section 2.1.2.4). By aligning with international standards and taking a modular, and interoperable approach, the SAF supports a flexible safety assurance process **suitable for multiple stakeholders** including regulators, consumer testing organisations, and manufacturers. This approach ensures compatibility with existing standards while providing the foundation for the creation of new standards (project objective #1). The inclusion of expert feedback via workshops (section 2.1.4.2) and the Midterm and Final Events (section 2.1.5.2, 2.1.5.3) further ensured relevance and adoption potential.

The **development of the SAF** has been closely linked with other tasks within the project, setting the blueprint for the development of its individual components. Work Package WP3 further developed the Environment block within the SAF, including concretisation and test instance allocation. Work Package WP4 developed a Harmonized V&V Simulation Framework for the Execute block. Work Package WP5 sets out the requirements and necessary building blocks for the Data Framework. Work Package WP6 developed and demonstrated the Data Framework, and Work Package WP7 demonstrated and validated the complete SAF based on various use cases. Work Package WP8 and WP9 engaged with international and European regulatory bodies, expert platforms, standardisation organisations, and twin projects (section 2.1.2, 2.1.3, 2.1.4, 2.1.5).

Key innovations include the formalisation of performance assurance blocks (chapter 4), integration with the Data Framework (chapter 5), and the creation of SAF Application Guidelines for both commercial and non-commercial users (chapter 7). The details illustrated around Safety Argument (section 4.3), and guidelines on applying them (section 7.3) provide the safety argumentation principles requested by project objective #2.

While the SAF presents a comprehensive approach to CCAM safety assurance, ongoing collaboration with regulators, consumer testers, and certification bodies remains essential to refine its application and scope. **Future developments** should focus on stress testing the SAF with wider applications such as in-service monitoring and reporting, AI, human machine interface (HMI) and V2X. Furthermore, ensuring scenario quality and completeness in external databases that connect to the SUNRISE Data Framework, should also be a key focus, as SUNRISE itself does not specifically address scenario creation. It is expected that follow-up projects (like SYNERGIES [71] and CERTAIN [72]) will explore these extensions and contribute to broader adoption of the SAF.

The SUNRISE SAF presented in this deliverable, reflects an effort to consolidate and structure the safety assurance process for CCAM systems in a way that can be applied, reviewed, and further refined by certifiers, regulators and industry. To those target users, it provides a **shared foundation for CCAM safety assurance**. Its adoption provides a clear path to consistent evaluation, faster certification, and stronger confidence in CCAM technologies. By embracing the SAF, stakeholders can **accelerate the safe and large-scale deployment of CCAM systems** on European public roads and thereby strengthen Europe's leadership in Cooperative, Connected and Automated Mobility technologies.

10 REFERENCES

[1] UNECE, "New Assessment/Test Method for Automated Driving (NATM)-Master Document," UNECE, Geneva, Switzerland, 2021.

[2] "ISO 26262-1:2018 Road vehicles — Functional safety," ISO, 2018.

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

[4] X. Sellart, J. Pont, A. Nogués, F. M. Navas, O. Opdencamp, A. Ballis, S. Dreher, J. Lorente, X. J. Zhang, G. Stettinger and D. Becker, "D2.1 Overview and gaps of existing safety assessment frameworks," 2023.

[5] "ISO 34503:2023 Road Vehicles — Test scenarios for automated driving systems — Specification for operational design domain," ISO, 2023.

[6] "EURO NCAP The European New Car Assessment Programme," [Online]. Available: https://www.euroncap.com/en. [Accessed 23 7 2025].

[7] "ISO: Global standards for trusted goods and services," [Online]. Available: https://www.iso.org/home.html. [Accessed 23 7 2025].

[8] "ASAM Association for Standardization of Automation and Measuring Systems," [Online]. Available: https://www.asam.net/standards/. [Accessed 23 7 2025].

[9] "The British Standards Institution," [Online]. Available: https://www.bsigroup.com/en-GB/products-and-services/standards/. [Accessed 23 7 2025].

[10] "The VVM project," [Online]. Available: https://www.vvm-projekt.de/en/. [Accessed 23 7 2025].

[11] "SAKURA RESEARCH PROJECT," [Online]. Available: https://www.sakura-prj.go.jp/project_info/. [Accessed 23 7 2025].

[12] "StreetWise: scenario based safety assessment for automated driving," [Online]. Available: https://www.tno.nl/en/digital/smart-traffic-transport/safety-assessment-automated-driving/streetwise/. [Accessed 23 7 2025].

[13] "Safety PoolTM scenario database," [Online]. Available: https://www.safetypool.ai/. [Accessed 23 7 2025].

[14] "ADScene, a scenario database for the validation of automated driving and driving-assistance systems via digital simulation," [Online]. Available: https://www.irt-systemx.fr/en/autonomous-vehicle-completion-of-the-experimental-programme-sam-and-launch-of-first-pilot-services/. [Accessed 23 7 2025].

[15] "Scenario Center," [Online]. Available: https://scenario.center/. [Accessed 23 7 2025].

[16] "AVL SCENIUS," [Online]. Available: https://www.avl.com/en-gb/testing-solutions/automated-and-connected-mobility-testing/avl-scenius. [Accessed 23 7 2025].

[17] X. Zhang, S. Khastgir, H. Asgari and P. Jennings, "Test Framework for Automatic Test Case Generation and Execution Aimed at Developing Trustworthy AVs from Both Verifiability and Certifiability Aspects," in *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2021.

[18] R. Donà, B. Ciuffo, A. Tsakalidis, L. Di Cesare, C. Sollima, M. Sangiogi and M. C. Galassi, "Recent Advancements in Automated Vehicle Certification: How the Experience from the Nuclear Sector Contributed to Making Them a Reality," *Energies,* vol. 15, 2022.

[19] "PEGASUS METHOD," [Online]. Available: https://www.pegasusprojekt.de/files/tmpl/Pegasus-Abschlussveranstaltung/PEGASUS-Gesamtmethode.pdf. [Accessed 23 7 2025].

[20] "HEADSTART," [Online]. Available: https://www.headstart-project.eu/. [Accessed 23 7 2025].

[21] "ArhitectECA2030," [Online]. Available: https://architect-eca2030.eu/. [Accessed 23 7 2025].

[22] "EU ADS Act," [Online]. Available: https://eur-lex.europa.eu/eli/reg_impl/2022/1426/oj/eng. [Accessed 23 7 2025].

[23] "UNECE - FRAV/VMAD Integration," [Online]. Available: https://wiki.unece.org/pages/viewpage.action?pageId=224002635. [Accessed 23 7 2025].

[24] "Assuring the safety of connected and autonomous vehicles," [Online]. Available: https://cp.catapult.org.uk/report/assuring-the-safety-of-connected-and-autonomous-vehicles/. [Accessed 23 7 2025].

[25] "ASAM OpenSCENARIO® XML," [Online]. Available: https://www.asam.net/standards/detail/openscenario-xml/. [Accessed 23 7 2025].

[26] "ASAM OpenDRIVE®," [Online]. Available: https://www.asam.net/standards/detail/opendrive/. [Accessed 23 7 2025].

[27] B. Bourauel, J. H. Vélez, S. v. Monfort, X. J. Zhang, J. Beckmann, E. Collado, P. Irvine, T. Menzel and M. D. Lio, "D8.1 Final report to vehicle safety bodies," 2025.

[28] "ISO 34502:2022 Road vehicles — Test scenarios for automated driving systems — Scenario based safety evaluation framework," [Online]. Available: https://www.iso.org/standard/78951.html. [Accessed 23 7 2025].

[29] "ISO 34504," [Online]. Available: https://www.iso.org/standard/78953.html. [Accessed 23 7 2025].

[30] "ASAM OpenODD®," [Online]. Available: https://www.asam.net/standards/detail/openodd/. [Accessed 23 7 2025].

[31] "ASAM OpenLABEL®," [Online]. Available: https://www.asam.net/standards/detail/openlabel/. [Accessed 23 7 2025].

[32] "ASAM OSI®," [Online]. Available: https://www.asam.net/standards/detail/osi/. [Accessed 23 7 2025].

[33] "Best Practices for Developing and Validating Simulations for Automated Driving Systems," [Online]. Available: https://www.sae.org/standards/content/j3279/. [Accessed 23 7 2025].

[34] "Taxonomy and Definition of Safety Principles for Automated Driving System (ADS)," [Online]. Available: https://www.sae.org/standards/content/j3206_202107/. [Accessed 23 7 2025].

[35] "PAS 1883:2025 Operational design domain (ODD) – Implementing BS ISO 34503," [Online]. Available: https://www.bsigroup.com/en-GB/insights-and-media/insights/brochures/operational-design-domain-implementation-of-bs-iso-34503/. [Accessed 23 7 2025].

[36] "BSI Flex 1891 Behaviour Taxonomy for Automated Driving System Applications," [Online]. Available: https://www.bsigroup.com/en-GB/insights-and-media/insights/brochures/bsi-flex-1891-behaviour-taxonomy-for-automated-driving-system-applications/. [Accessed 23 7 2025].

[37] "BSI Flex 1889 - Natural Language Description for Automated Driving Systems," [Online]. Available: https://www.bsigroup.com/en-GB/insights-and-media/insights/brochures/bsi-flex-1889-natural-language-description-for-automated-driving-systems/. [Accessed 23 7 2025].

[38] "Self-assessment, protection & healing tools for a resilient CCAM’s ecosystem," [Online]. Available: https://selfy-project.eu/. [Accessed 22 7 2025].

[39] "Integrated 4D driver modelling under uncertainty," [Online]. Available: https://i4driving.eu/. [Accessed 22 7 2025].

[40] "SINFONICA," [Online]. Available: https://sinfonica.eu/. [Accessed 23 7 2025].

[41] "Partners for Automated Vehicle Education," [Online]. Available: https://pavecampaign.org/uk/. [Accessed 23 7 2025].

[42] "OPTIPEX," [Online]. Available: https://optipex.eu/. [Accessed 23 7 2025].

[43] "CONNECT," [Online]. Available: https://horizon-connect.eu/. [Accessed 23 7 2025].

[44] D. Ozmendji, J. Pont , S. de Vries , X. J. Zhang , A. Arrieta , J. Beckmann and M. N. Doncel , "D9.3 Final external cooperation report," 2025.

[45] "AVSC," [Online]. Available: https://avsc.sae-itc.com/.

[46] "KATRI," [Online]. Available: https://main.kotsa.or.kr/eng/road/safetyResearch.do?menuCode=03020000.

[47] "Transport Canada," [Online]. Available: https://tc.canada.ca/.

[48] "SET Level," [Online]. Available: https://setlevel.de/en/.

[49] "DIVP," [Online]. Available: https://divp.net/.

[50] "ASAM TestSpecification," [Online]. Available: https://www.asam.net/standards/asam-test-specification/.

[51] "FMI," [Online]. Available: https://fmi-standard.org/.

[52] "SAE J3016," [Online]. Available: https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic.

[53] "UN Regulation No. 157 - Automated Lane Keeping Systems (ALKS)," [Online]. Available: https://unece.org/transport/documents/2021/03/standards/un-regulation-no-157-automated-lane-keeping-systems-alks.

[54] A. Thorsén, O. Bartels, C. Berger, T. Bouraffa, C. Kaya, M. Muro, A. Bolovinou, I. Panangiotopoulos, D. Becker, A. Farooqui, M. Skoglund, P. Stålberg, F. Warg, E. Arnoux, O. Op den Camp, G. Ben Nejma, A. Bruto da Costa and P. Irvine, "D3.1: Report on baseline analysis of existing methodology," SUNRISE, 2023.

[55] E. de Gelder, J.-P. Paardekooper, A. Khabbaz, H. Elrofai, O. Op den Camp, S. Kraines, J. Ploeg and B. de Schutter, "Towards an Ontology for Scenario Definition for the Assessment of Automated Vehicles: An Object-Oriented Framework," *IEEE Transactions on Intelligent Vehicles,* pp. 1-1, 2022.

[56] E. Stellet, M. R. Zofka, J. Schumacher, T. Schamm, F. Niewels and J. M. Zöllner, “Testing of Advanced Driver Assistance Towards Automated Driving: A Survey and Taxonomy on Existing Approaches and Open Questions,” in *IEEE 18th International Conference on Intelligent Transportation Systems*, 2015.

[57] E. de Gelder, O. Op den Camp and N. de Boer, “Scenario categories for the assessment of automated vehicles,” 2020. [Online]. Available: http://www.cetran.sg/publications.

[58] E. de Gelder, J.-P. Paardekooper, O. Op den Camp and B. De Schutter, “Safety assessment of automated vehicles: how to determine whether we have collected enough field data?,” *Traffic Injury Prevention,* vol. 20, pp. S162-S170, 2019.

[59] J. Beckmann , X. J. Zhang , E. de Gelder, E. van Hassel , M. Nieto , G. Ben Nejma and M. Grabowski , "D5.2 Harmonised descriptions for content of CCAM safety assessment data framework," 2025.

[60] S. Chen, S. Khastgir, B. Islam and J. Paul, "Identifying Accident Causes of Driver-Vehicle Interactions Using System Theoretic Process Analysis (STPA)," 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2020.

[61] C. C. e. al., "MetaScenario: A Framework for Driving Scenario Data Description, Storage and Indexing," in *IEEE Transactions on Intelligent Vehicles*, 2023.

[62] "VVM Project," [Online]. Available: https://www.vvm-projekt.de/en/.

[63] "UN Regulation No. 157 on uniform provisions concerning the approval of vehicles with regards to Automated Lane Keeping System," UN Economic Commission for Europe,

Inland Transport Committee, World Forum for Harmonization of Vehicle Regulations, Geneva, 2021.

[64] E. de Gelder and O. Op den Camp, "Procedure for the Safety Assessment of an Autonomous Vehicle using Real-World Scenarios," in *FISITA 2020 World Congress*, Prague, 2020.

[65] N. Kauffmann, F. Fahrenkrog, L. Drees and F. Raisch, "Positive risk balance: a comprehensive framework to ensure vehicle safety," in *Ethics and Information Technology*, 2022.

[66] P. Koopman, "How Safe is Safe Enough? Measuring and Predicting Autonomous Vehicle Safety," in *Carnegie Mellon University*, Pittsburgh, 2022.

[67] E. de Gelder and O. Op den Camp, "How certain are we that our automated driving system is safe?," in *Traffic Injury Prevention, Special issue for the 27th International Technical Conference on the Enhanced Safety of Vehicles (ESV)*, 2023.

[68] S. U. e. al, "Data Sources for Baseline Generation – Overview, Grading, and Recommendations," Deliverable D4.1 of the Horizon Europe project V4SAFETY, 2024.

[69] F. e. a. Fahrenkrog, "Prospective Safety Assessment Framework – Instruction," Deliverable D2.1 of the Horizon Europe project V4SAFETY, 2024.

[70] "V4SAFETY," [Online]. Available: https://v4safetyproject.eu/. [Accessed 23 7 2025].

[71] "SYNERGIES," [Online]. Available: https://synergies-ccam.eu/.

[72] "CERTAIN," [Online]. Available: https://cordis.europa.eu/project/id/101203230.

[73] ISO/TC 22/SC 36/WG 7, "ISO/TR 21934-1:2021 Road vehicles — Prospective safety performance assessment of pre-crash technology by virtual simulation — Part 1: State-of-the-art and general method overview," ISO, 2021-06.

[74] ISO 21934, "Road vehicles — Prospective safety performance assessment of pre-crash technology by virtual simulation — Part 2: Guidelines for application," International Organization for Standardization, expected 2023.