D5.1

Requirements for CCAM safety assessment data framework content

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	WP5: Content harmonisation of scenario data framework
Deliverable number	D5.1: Requirements for CCAM safety assessment data framework content
Deliverable responsible	Fadi Alakkad (University of Warwick)
Status - Version	Final – V1.1
Submission date	15/05/2024
Keywords	Requirements, Data Framework, Scenario Database, CCAM

Authors

Name	Organisation
Fadi Alakkad	WMG, University of Warwick (UoW)
Xizhe Xhang	WMG, University of Warwick (UoW)
Siddartha Khastgir	WMG, University of Warwick (UoW)
Erwin de Gelder	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO (TNO)
Sjef van Montfort	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO (TNO)
Jorge Lorente Mallada	TOYOTA MOTOR EUROPE NV (TME)
Edwin van Hassel	SIEMENS INDUSTRY SOFTWARE NETHERLANDS BV (SISW)
Maximilian Grabowski	BUNDESANSTALT FUER STRASSENWESEN (BASt)
Jobst Beckmann	RHEINISCH-WESTFAELISCHE TECHNISCHE HOCHSCHULE AACHEN (ika)
Xavier Boabén	Applus IDIADA
Aleix Nogués	Applus IDIADA
Sergi Vidal	Applus IDIADA

Thaddaeus Menzel	IDIADA Fahrzeugtechnik GmbH
Ilias Panagiotopoulos	INSTITUTE OF COMMUNICATION AND COMPUTER SYSTEMS (ICCS)
Athanasios Ballis	INSTITUTE OF COMMUNICATION AND COMPUTER SYSTEMS (ICCS)

Quality Control

	Name	Organisation	Date
Peer review 1	Olaf Op den Camp, Sjef van Montfort	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO (TNO)	06/07/2023
Peer review 2	Ilias Panagiotopoulos, Athanasios Ballis	INSTITUTE OF COMMUNICATION AND COMPUTER SYSTEMS (ICCS)	11/07/2023
Peer review 3	Stefan de Vries	IDIADA	19/07/2023
Peer review 4	Marcos Nieto Doncel	VICOMTECH	21/07/2023

Version history

Version	Date	Author	Summary of changes
0.1	02/05/2023	Fadi Alakkad	First draft of document structure.
0.2	08/06/2023	Fadi Alakkad, Erwin De Gelder, Sjef van Montfort, Mohsen Alirezaei, Edwin van Hassel, Maximilian Grabowski, Thaddaeus Menzel, Ilias Panagiotopoulos, Athanasios Ballis Jorge Lorente Mallada, Jobst Beckman,Xavier Boabén, Aleix Nogués, Sergi Vidal	The first draft of the content after listing the requirements and clusters from the excel sheet.

0.3	03/07/2023	Fadi Alakkad, Erwin De Gelder, Sjef van Montfort, Mohsen Alirezaei, Edwin van Hassel, Maximilian Grabowski, Thaddaeus Menzel, Ilias Panagiotopoulos, Athanasios Ballis Jorge Lorente Mallada, Jobst Beckman,Xavier Boabén, Aleix Nogués, Sergi Vidal, Siddartha Khastgir, Xizhe Zhang	Final draft version.
0.4	21/07/2023	Olaf Op den Camp, Sjef van Montfort, Ilias Panagiotopoulos, Athanasios Ballis, Stefan de Vries, Marcos Nieto Doncel	Peer review.
0.5	08/08/2023	Fadi Alakkad, Siddartha Khastgir, Xizhe Zhang	Revise the document in accordance with the feedback provided by the reviewers.
0.6	09/08/2023	Fadi Alakkad, Erwin De Gelder, Sjef van Montfort, Mohsen Alirezaei, Edwin van Hassel, Maximilian Grabowski, Thaddaeus Menzel, Ilias Panagiotopoulos, Athanasios Ballis Jorge Lorente Mallada, Jobst Beckman,Xavier Boabén, Aleix Nogués, Sergi Vidal, Siddartha Khastgir, Xizhe Zhang	1st revision of the requirements list.
0.6.1	15/08/2023	Fadi Alakkad, Erwin De Gelder, Sjef van Montfort, Mohsen Alirezaei, Edwin van Hassel, Maximilian Grabowski, Thaddaeus Menzel, Ilias Panagiotopoulos, Athanasios Ballis Jorge Lorente Mallada, Jobst	2nd revision of the requirements list.

		Beckman,Xavier Boabén, Aleix Nogués, Sergi Vidal, Siddartha Khastgir, Xizhe Zhang	
0.6.2	17/08/2023	Fadi Alakkad, Erwin De Gelder, Sjef van Montfort, Mohsen Alirezaei, Edwin van Hassel, Maximilian Grabowski, Thaddaeus Menzel, Ilias Panagiotopoulos, Athanasios Ballis Jorge Lorente Mallada, Jobst Beckman,Xavier Boabén, Aleix Nogués, Sergi Vidal, Siddartha Khastgir, Xizhe Zhang	3rd revision of the requirements list.
0.7	31/08/2023	Fadi Alakkad, Siddartha Khastgir, Xizhe Zhang	Revision of the document.
1.0	04/09/2023	Fadi Alakkad, Siddartha Khastgir, Xizhe Zhang	Final submitted version.
1.1	15/05/2024	Fadi Alakkad, Siddartha Khastgir	Updates resulting from EC review

Legal disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATION AND ACRONYMS

Abbreviation	Meaning
ADAS	Advanced Driver Assistance System
ADS	Automated Driving System
ALKS	Automated Lane Keeping System
API	Application Programming Interface
CCAM	Connected, Cooperative and Automated Mobility
CL	Cluster
D	Deliverable
EC	European Commission
EU	European Union
FRAV	UNECE Working Group on Functional Requirements for Automated and Autonomous Vehicles
GDPR	General Data Protection Regulation
GRVA	Working Group on Automated and Connected Vehicles
GUI	Graphical User Interface
JRC	Joint Research Centre
MVWG	Motor Vehicle Working Group
NCAP	New Car Assessment Programme
OD	Operational Domain
ODD	Operational Design Domain
OEM	Original Equipment Manufacturer
R&D	Research and Development
RE	Requirement
REST-API	Representational State Transfer API
SAF	Safety Assurance Framework
SCDB(s)	Scenario Database(s)
SuT	System under Test
THW	Time Headway
TTC	Time To Collision
UI	User Interface
UN	United Nations
UNECE	United Nations Economic Commission for Europe
V2X	Vehicle to Everything
VMAD	UNECE Working Group on Validation Method for Automated Driving
WP	Work Package

EXECUTIVE SUMMARY

The SUNRISE project aims to accelerate the safe deployment of Cooperative, Connected, and Automated Mobility (CCAM) technologies by developing and demonstrating a Safety Assurance Framework (SAF). The project addresses the challenges of safety assurance in CCAM systems by focusing on a mixture of physical and virtual testing, scenario databases, validation methods, and harmonisation of standards.

To develop the Safety Assurance Framework, one part of the project identifies requirements from different stakeholders, including regulatory bodies, consumer testing organisations, solution providers, and research institutes. The requirements cover aspects such as scenario description, scenario content, and data framework. The project emphasises the importance of a common approach to testing and validation, enabling standardised results, improved analysis, and comparability for the introduction of CCAM systems.

This deliverable presents the requirements identified by stakeholders for the data framework and scenario databases. It describes the process followed to gather and agree on the requirements, as well as the clustering of requirements. The intended audience for the deliverable includes consortium partners working on technical tasks 5.2, 5.3, 6.1 and 6.2.

The deliverable also provides a short description of the data framework and the external scenario databases content. The data framework enables user access to these external scenario databases, including uploading new scenarios, searching for scenarios, and storing/retrieving test results. The external scenario databases content encompasses the format for storing scenarios, metadata, relevant outputs, scenario data origin, and the scenarios themselves.

The SUNRISE T5.1 follows a collaborative approach to gather requirements, involving partners' expertise and extensive discussions. The requirements are refined through iterations, ensuring consistency, coherence, and alignment with the SUNRISE's vision. The finalised requirements are grouped into clusters based on their topics, and each cluster is assigned a unique identifier for easy reference.

The deliverable serves as a fundamental source for various work packages within the SUNRISE project and will be used in tasks related to data framework harmonisation, quality metrics development, interface synchronisation, and architecture development.

Overall, the SUNRISE project strives to establish a comprehensive Safety Assurance Framework and a federated scenario databases approach that will enhance the validation and assurance of CCAM systems, promote harmonisation, and facilitate the safe deployment of these technologies.

1 INTRODUCTION

1.1 SUNRISE project

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.

CCAM systems need to demonstrate reliability in all driving scenarios, requiring robust safety argumentation. It is already 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 in this mixture for cost efficiency reasons.

Several worldwide initiatives have started to develop test and assessment methods for automated driving functions. These initiatives have already moved 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 fact that there is a 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 whole 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 are another issue tackled by several initiatives and projects, providing silo solutions. A clear concrete approach should be used (at least at the 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 in order 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 initiatives as a baseline, it becomes necessary to move to the next level in the concrete specification and demonstration of a commonly accepted **Safety Assurance Framework** (**SAF**) for the safety validation of CCAM systems, including a broad portfolio of use cases 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 Safety Assurance Framework is the main element to be developed in the SUNRISE project. As the following figure 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 by creating and sharing a European federated database framework centralising detailed scenarios for testing of CCAM functions and systems in a multitude of relevant test cases, based on a harmonised simulation and test environment with standardised, open interfaces and quality-controlled data exchange.

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.

The following figure shows the general overview of the SUNRISE project.

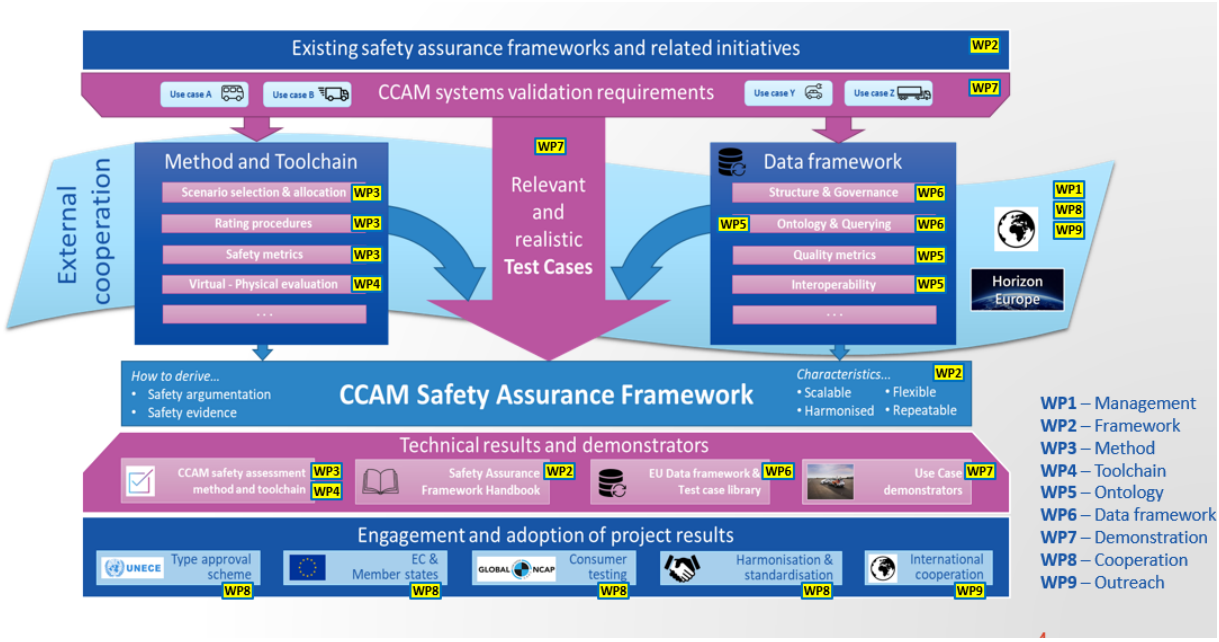

Figure 2: Overview of the SUNRISE Project

1.2 Purpose of the deliverable

The purpose of this deliverable is to present and describe the requirements identified by stakeholders for a **SUNRISE Data Framework** which includes the federated scenario databases. The stakeholders involved in this process include Homologation Technical Service (e.g., UNICE VMAD, FRAV, GRVA, WP.29), consumer testing organisations (e.g., Euro NCAP), Original Equipment Manufacturers (OEMs), scenario database host, and research institutes, having each of them different use cases related the data framework and SCDB content (SUNRISE D2.1).

The deliverable provides an overview of the data framework and scenario database content definition (see chapter 2). The data framework serves as the backbone for user access to the federated scenario databases, facilitating activities such as uploading new scenarios, searching for scenarios, and storing or retrieving test results. The federated scenario databases content includes the format for storing scenarios, metadata, relevant outputs, scenario data origin, and the scenarios themselves. It should also include different data sources (accidents, naturalistic driving studies, field operational tests, etc.) based on their feasibility and added value.

Additionally, the document outlines different methodologies to gather and agree on the requirements, highlighting the collaborative nature of the effort (see section 3.2). It also includes information on how the requirements have been clustered based on their similar characteristics or attributes (see section 3.3), making it easier for stakeholders to locate specific requirements within the document. The aim is to ensure that the requirements are comprehensive, cover all relevant data sources used in the data framework, and contribute to or define harmonized descriptions in consensus with all stakeholders and the constituting scenario databases. This is crucial to ensure all the scenario databases included under the

SUNRISE federated layer provide same level of coverage of the required properties and features and encompass the full spectrum.

Furthermore, the deliverable serves as a fundamental source for various work packages within the SUNRISE project, such as WP7 about use cases and framework demonstration instances development, WP8 on cooperation with international vehicle safety bodies & WP9 on dissemination and international cooperation. It will be utilised in tasks related to data framework harmonisation, quality metrics development, interface synchronisation, and architecture development.

In summary, the primary purpose of this deliverable is to provide a detailed and comprehensive description of the requirements for the data framework and the constituting federated scenario databases. Additionally, it serves as a reference for related tasks (see section 1.4) within the SUNRISE project, promoting harmonisation and consensus among stakeholders.

1.3 Intended audience

The internal project audience is the key stakeholder of this deliverable as its results serve as a fundamental source for various work packages of SUNRISE. It contains requirements for the SUNRISE data framework from many different stakeholders. Task 5.2 is directly dependent on these results. It harmonizes the data framework and the scenario database (SCDB) content based on the developed requirements. This harmonisation is done mostly via templates, formats, standards, technical specifications, and especially ontologies. Results from this harmonisation support Task 5.3, where quality metrics are developed for SCDB content based on the aforementioned ontologies; Task 4.2, which synchronizes all data framework related topics to ensure correct interfaces; and Task 4.4, that develops the harmonized interfaces between core subsystems, including the integrated scenario databases. WP6 aims to create an architecture for the data framework, where the interfaces follow the results of WP5. Furthermore, the developed data framework will be considered in a data management plan to ensure proper data treatment. Furthermore, Task 6.1 will utilize the specified requirements to formulate input standards for the SCDBs. Correspondingly, Task 6.2 will be dedicated to crafting output standards from the SCDBs, encompassing query formats, criteria, and the linkage between the SCDBs and the data framework. Additionally, Task 6.3 will delineate processes for integrating SCDBs into testing and validating operations. To validate the content scenarios originating from the SCDBs for end users, the requirements outlined in D5.1 will be valuable.

Another key stakeholder are the SUNRISE project external ones and the development of its federated approach, and the integration of different scenario databases are vital for type approval authorities and vehicles manufacturers. Results of this project will help in the homologation process of Advanced Driver Assistance System (ADAS) and Automated Driving System (ADS) functions to ensure safety and usability in many different countries. The intended audience on a regulatory level are different working groups at the United Nations Economic Commission for Europe (UNECE) and the European Commission (EC) with its member states. Their working groups usually are composed of contracting parties, national

type-approval authorities (e.g., KBA, RDW), industry associations (e.g., OICA, CLEPA), technical services (e.g., TÜV, DEKRA), and several non-profit organizations (e.g., FIA, ETSC).

For UNECE the relevant informal working groups are the "*Validation Method for Automated Driving* (VMAD)" and the "*Functional Requirements for Automated and Autonomous Vehicles* (FRAV)" as part of the *Working Group on Automated and Connected Vehicles* (GRVA). These groups are responsible for the development of guidelines with regard to requirements and validation methods on automated functions and vehicles, which later will be adapted to UN Regulations of UN member states.

For the European Commission an important working group is the *Motor Vehicle Working Group (MVWG)* and the *Joint Research Centre (JRC)* as a research institute for the EC. The EC is also able to define regulations that affect its member states.

Moreover, Euro NCAP and its working groups are part of the intended audience as they perform highly regarded consumer testing(s) for all kinds of safety aspects of vehicles. Assisted and automated driving must be tested and rated accordingly, which can be supported by the SUNRISE data framework. The cooperation of SUNRISE with standardisation bodies (ASAM, ISO and SAE) is vital to foster new standardisation activities that are based on SUNRISE' results.

From an industrial perspective, vehicle manufacturers are the main audience of this project as the results will help them to ensure proper verification and validation methods for the safety of ADAS and ADS functions.

The following figure demonstrates the workflow of SUNRISE Work Package 5, and the links between various tasks and this deliverable (output of T5.1)

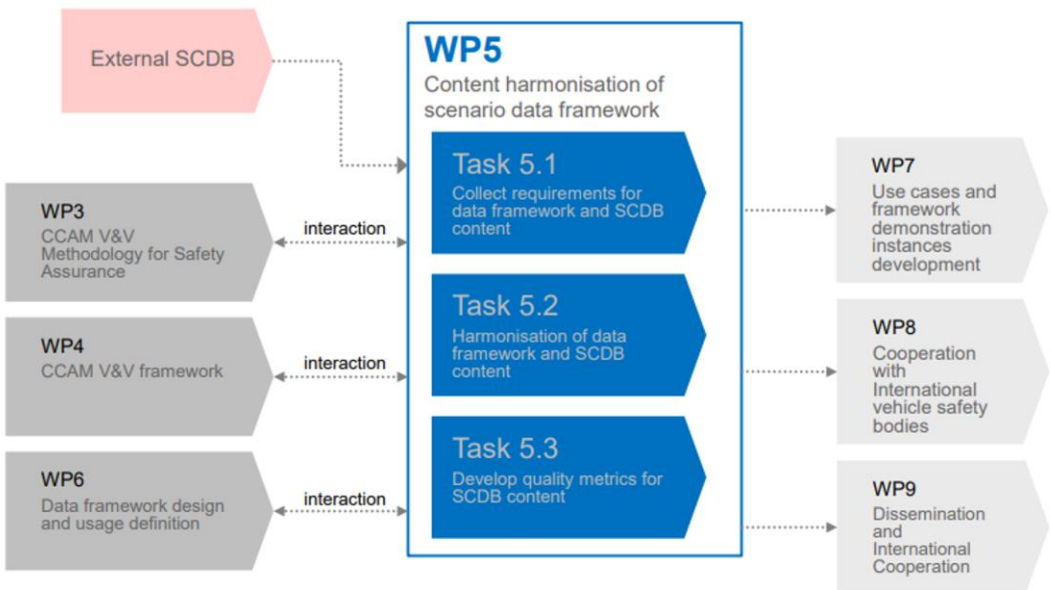

Figure 3: Workflow of SUNRISE Work Package 5.

1.4 Deliverable structure and relation with other work packages

The content of this deliverable is divided in the following chapters:

Chapter 2: **SUNRISE Data framework**. This chapter presents the data framework and the European federated scenario databases (SCDBs), and their meaning within the scope of SUNRISE Safety Assurance Framework.

Chapter 3: **Requirements.** The method followed to gather the requirements for the data framework as well as for the Scenario Database content is described. Additionally, this section describes the clusters defined to group the identified requirements. As main outcome, this chapter contains the list of requirements grouped by clusters.

Chapter 4: **Conclusions.** It describes the main highlights of the obtained results and how they will be used in the project.

The content of D5.1 has been aligned with D2.1. Furthermore, D5.1 output will be used mainly by Task 5.2, where the list of derived requirements will be used as input for this task that will deal with harmonisation of the data framework content. The contribution to Task 5.2 is highly relevant since this task has a direct relation to WP3, WP4 and WP6. Further details on the relation between this deliverable and other SUNRISE work packages and tasks can be seen under Section 1.3

2 SUNRISE DATA FRAMEWORK

2.1 Introduction

The objective of this chapter is to present a thorough comprehension of two fundamental concepts pertaining to databases for cooperative, connected and automated mobility (CCAM) safety assessment: **the data framework** and the **European federated scenario databases (SCDBs) content definition**. Subsequently, pertinent, and coherent requirements for the scenario data framework can be established, building upon a shared understanding as the foundation. The knowledge of these terms was acquired through partner knowledge, involving extensive collaboration, discussions, and feedback throughout this task.

The harmonized data framework described in this chapter will bring significant benefits to the various stakeholders involved in the SUNRISE project, as defined in section 1.3 of this deliverable. Additionally, it will enable the use of participating databases in a federated approach. With the implementation of a harmonized data framework, stakeholders will have the ability to access scenarios from the diverse range of existing databases connected to it.

The following sections provide a more detailed overview of both the data framework, as well as the federated scenario databases content.

2.2 Data framework

The data framework is a set of tools to enable user access to the various scenario databases that connect to it through a federated interface. This primarily concerns searching for scenarios in the various external scenario databases. Next to this, the data framework may also include facilitating the storing and retrieving test results of specific test cases in relevant scenario database or in a separate database alongside the existing scenario databases. All relevant information, especially descriptions / metadata, from the scenarios must be accessible and searchable through the data framework through a common and agreed query mechanism.

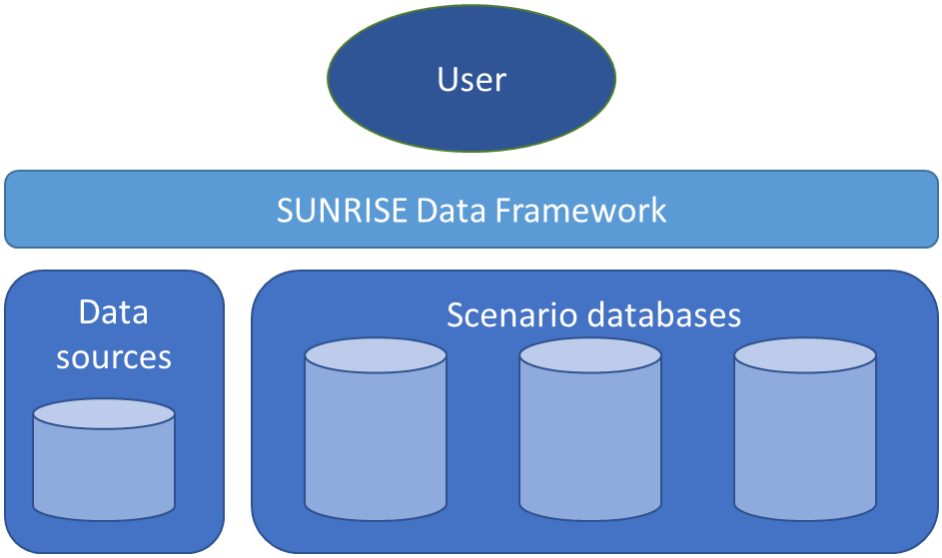

Figure 4: Environment of SUNRISE Data Framework

Figure 4 provides an overview of the envisioned data framework. This diagram is used to indicate where the data framework sits. In this framework, the user accesses the scenario databases through the data framework. The user is diverse can consist of many roles within the user organisations. The data framework will retrieve the scenarios from the external scenario databases used in the Sunrise project. However, before retrieving the scenarios, certain processing steps are necessary to harmonize the scenario data which will be carried out by the individual databases. Furthermore, Various Data sources may be used in addition to the scenario databases to enrich the retrieved results.

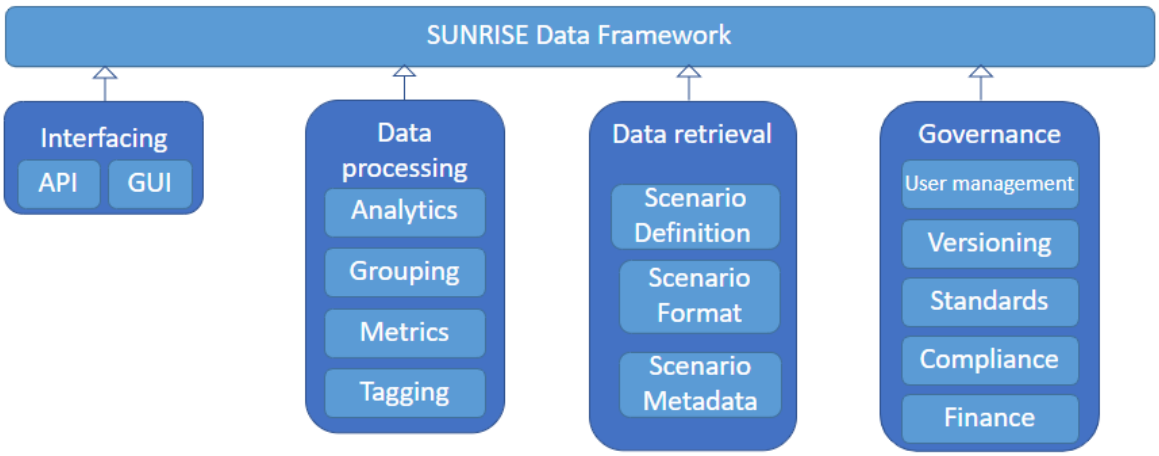

Figure 5: Concept of SUNRISE Data Framework (class diagram)

The Figure 5 demonstrates various components that are aggregated in the data framework. The interfacing is the process of allowing users to request scenarios from the data framework through the GUI and it is the interaction point between the data framework and the external SCDBs via the API. The data processing function handles the scenario grouping and scenario tags based on agreed taxonomy. Additionally, the data framework allows for the retrieval of scenario data from external scenario databases. Users can query the framework, which then

outputs relevant scenarios based on metadata information or scenario labels. This enables efficient searching and retrieval of specific scenarios. Finally, Governance defines the process for maintaining and improving the data framework. Furthermore, once tests have been executed using the scenario data, the results can be conveniently stored into a test results database as identified by the user of the data framework.

The federated layer plus any connected databases is integral part of the SAF, because the execution and the analysis sections as defined in the draft SAF rely on the scenarios to be retrievable from the federated layer. Below is a figure that explains how data framework enables SAF.

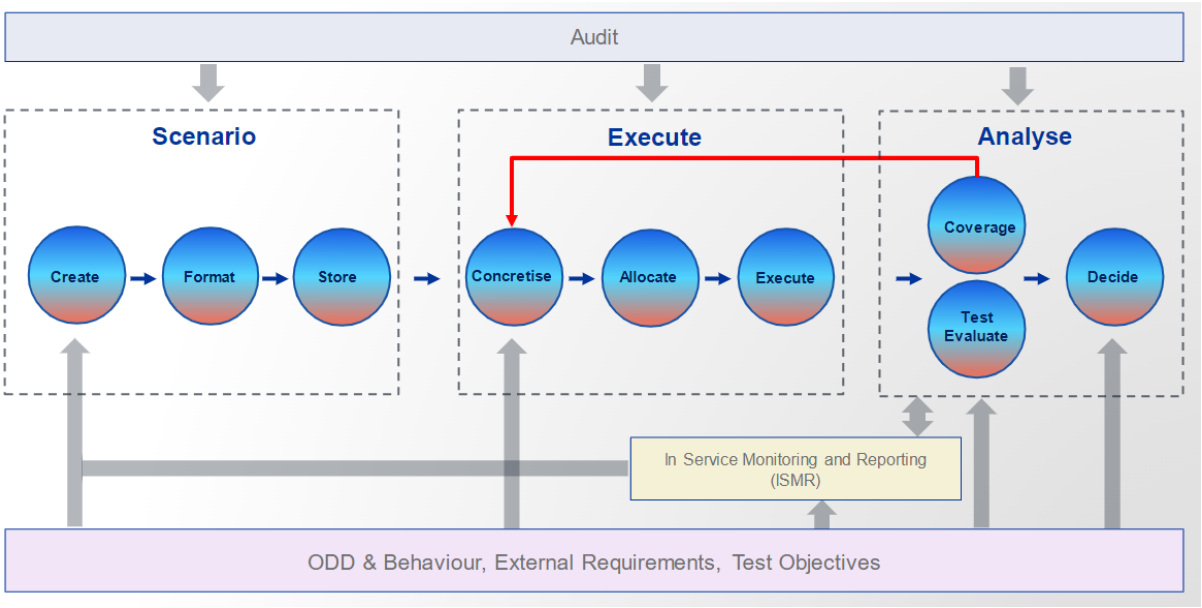

Figure 6: Safety Assurance Framework (Draft)

As shown on the figure above, the data framework will bridge the Scenarios and Execute stages, and serves as a crucial link, seamlessly connecting SCDBs through federated layers to the broader testing workflow. Leveraging a common and standardized data framework, alongside a defined set of requirements outlined in this deliverable, enables the SAF to effectively access and retrieve scenarios stored within SCDBs.

Overall, the SUNRISE data framework facilitates the retrieval, processing, and integration of scenario data within the associated scenario databases, enabling streamlined data and efficient utilisation of scenarios for testing purposes.

2.3 Scenario database content

When discussing the contents of a scenario database, it is important to differentiate between the scenario database format and the actual scenario database content stored within various databases. The scenario database format refers to the structure and organization of scenarios, which encompasses aspects such as parameter lists, file formats, metadata types,

and more. Metadata plays a crucial role in defining scenarios by providing additional information that is not inherently part of the scenario content. This includes relevant outputs and associated criteria, the source of the scenario data (e.g., expert-created data or data from a specific collection campaign), as well as test specifications and plans.

On the other hand, the scenario database content pertains to the actual scenarios themselves, as defined according to the scenario database format. These scenarios can consist of diverse data, including information about road networks, environments, weather conditions, dynamic objects, and more. Different levels of scenarios exist, as defined in Sunrise deliverable D3.1 ([1]), which uses the definition from [2], as shown in Figure 6. A scenario database could contain concrete, logical, and possible abstract scenarios. Functional scenarios will not be part of a scenario database as these are not machine-readable.

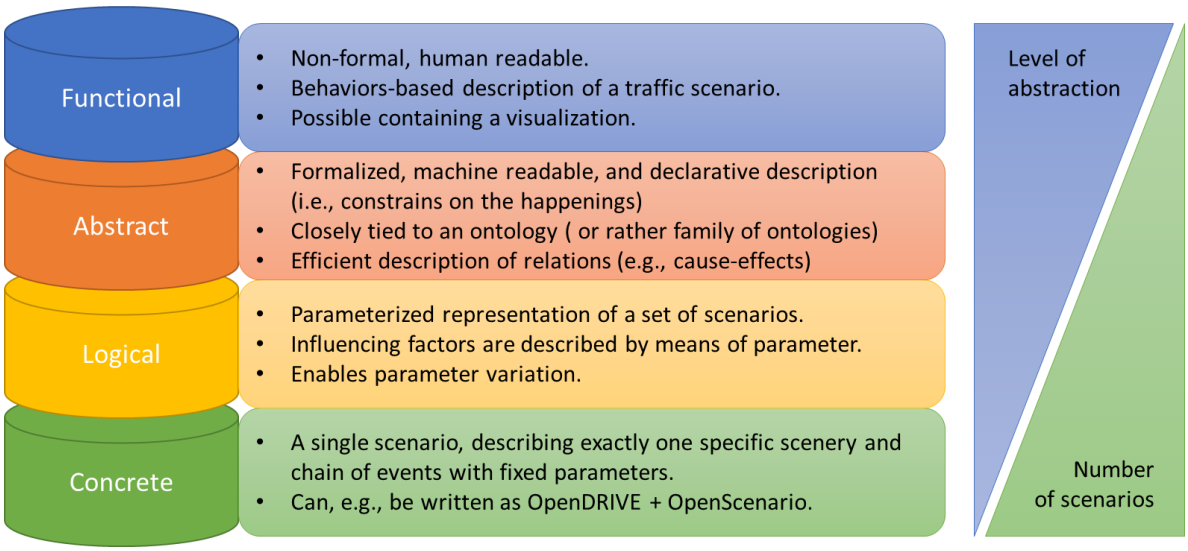

Figure 7: The relationship between functional, abstract, logical and concrete scenarios ([1]).

3 REQUIREMENTS

3.1 Introduction

The collection of requirements in task T5.1 of the SUNRISE project is a collaborative effort which aims to produce requirements for the data framework. The next sections provide details on the process followed to create and agree on the list of requirements. In general terms, the requirements have been gathered as follows:

1. Each partner proposed a list of requirements that are relevant for their context and objectives. The requirements should be clear, concise, and measurable, and include a definition and a rationale.

2. The requirements are shared with all the partners through a shared document, where they can be reviewed, and edited online. Partners can provide feedback and suggest changes. The goal is to ensure consistency, coherence, and alignment with the project vision.

3. Based on the feedback, the requirements are revised and refined during various T5.1 meetings. The process iterates until all the partners agree on the requirements. This may involve merging, splitting, or removing some requirements, and clarifying their definitions and scope.

4. Once all the requirements have been finalised, they are grouped into clusters according to their topic or theme. For example, a cluster could be related to User Management, API, Interfacing or UI/GUI. The clusters are also shared and debated among the partners, to avoid duplication or contradiction.

5. Each cluster is assigned an unambiguous description, and a unique identifier (e.g., CL1) to simplify requirements utilisation in further development stages of the project.

3.2 Methodologies used to gather the requirements

All partners involved in this task have elaborated a list of relevant requirements. With a heterogeneous group of partners going from academia or research background up to industrial partners, a wide variety of partners experiences has been achieved to the topic of the use of scenario data for safety assurance of CCAM.

This has allowed to bring in requirements discussed also at relevant projects in individual countries, such as PEGASUS project family [3] in Germany, SAKURA [4] in Japan, or in initiatives such as Streetwise [5], Safety Pool [6], ADScene [7], or ArchitectECA2030 [8]

Besides that, some partners have expertise on specific topics, such as connectivity, cybersecurity, or type approval, which has helped to elaborate concrete requirements considering those.

Furthermore, during the survey activity involving stakeholders in the ecosystem, more than 50 organizations were contacted at significant international events, thereby ensuring a thorough and precise enhancement of these requirements.

Moreover, external stakeholder engagement has primarily occurred through standardization efforts, collaborative projects, dissemination activities, and participation in national/international events. These stakeholders have provided feedback in informal manners, predominantly through verbal exchanges, which has been assimilated into the requirement table. It is important to note that some feedback might have been confidential or expressed as personal opinions not necessarily representing the views of their respective organisations. Due to these reasons, formal recording of such feedback may not have been feasible, yet it remains integral to the project progress.

Acknowledging the potential challenge of traceability due to the informal nature of this feedback channel, the information remains integral to the project's progress.

The descriptions of the methodologies used to gather the requirements are listed in Annex 1.

3.3 Requirement clusters

To help identify and group requirements that share similar characteristics or attributes, the partners have grouped the requirements into clusters. This enables a better understanding of the requirements by highlighting the relationships and patterns among them. The requirements are being collected in the following clusters.

Table 1: Clusters definitions

Nr.	Cluster	Description
CL1	User Management	This cluster covers handling user access, managing user accounts, defining roles, and identifying the main users.
CL2	Interfacing	Refers to requirements related to the interaction and exchange of data between the data framework and external scenario databases.
CL3	UI/GUI	Involves requirements related to the visual representation of data and user interaction with the data framework
CL4	Database Data Processing/Analytics	Involves requirements related to functionalities to process, analyse, and manipulate data within the data framework.
CL5	Version Control	Involves requirements related to tracking and managing the versions of scenarios within the data framework.
CL6	File Attachment	Involves requirements related to associate and attach different types of files to scenarios within data framework.
CL7	Structure of Scenario Representation	Involves requirements related to organising and categorising scenarios data framework
CL8	Standards Alignment	Involves requirements related to align the input and output formats with relevant standards.
CL9	Scenario Filtering/Searching	Refers to a set of requirements or functionalities related to searching and filtering scenarios.
CL10	Legal and Compliance	Refers to requirements related to the compliance with legal requirements, and the protection of legal rights and obligations.

CL11	Scenario Metadata Association	Involves requirements that focuses on capturing and associating essential metadata elements with scenarios.
CL12	Scenario Metrics	Involves requirements related to measuring the quality of scenarios within the data framework.
CL13	Scenario Definition and Format	Involves requirements related to the format and structure of scenarios within the data framework.
CL14	Scenario Data Source Categorisation	Involves requirements of the data sources that the data framework.
CL15	Scenario Tagging/Labelling	Involves requirements related to tagging and labelling scenarios within the data framework.
CL16	Test Results	involves requirements related to test results in the data framework.

3.4 Requirement list

In this section, the requirements are presented based on the corresponding clusters:

CL1 - User Management Requirements:

Table 2: User Management Requirements.

Nr.	Requirements	Description
RE1	The data system shall allow to create, read, update and delete user roles and profiles.	The system shall have capabilities to administrate roles of users, with a user profile sheet that includes user details, access rights, history of actions, etc.
RE2	The data system shall apply user rights to scenarios and other content as defined in the underlying scenario databases.	User shall only be able to access its licensed scenarios and databases. Not everybody has the same rights.
RE3	The data shall allow to record, inspect and delete user history.	For quality control, it is important to data framework. Users also have a right to be forgotten, hence the option to delete user history.

CL2 – Interfacing Requirements:

Table 3: Interfacing Requirements

Nr.	Requirements	Description
RE4	The Data Framework shall allow querying of the various underlying SCDBs for retrieval of scenario data.	The Data Framework shall be able to communicate with different SCDBs to retrieve scenario data when needed.
RE5	The API of the Data Framework shall be designed so that implementation at consumer side takes less than a week of work.	The API of the Data Framework shall be easy to use and understandable to the software engineers linking to the Data Framework.

RE6	The API shall be based on REST-API.	The API should be based on state-of-the-art and standards (e.g., REST-API).

CL3– UI/GUI Requirements:

Table 4: UI/GUI Requirements

Nr.	Requirements	Description
RE7	The Data Framework shall provide a graphical user interface (GUI) that allows searching of scenarios.	A user interface helps the user to navigate through the functionalities of the Data Framework, such as searching for relevant scenarios, or exporting them into specific formats.
RE8	The Data Framework shall provide a GUI that allows exporting scenarios.	A user interface helps the user to navigate through the functionalities of the Data Framework, such as searching for relevant scenarios, or exporting them into specific formats.

CL4– Database Data Processing/Analytics Requirements:

Table 5: Database Data Processing/Analytics Requirements

Nr.	Requirements	Description
RE9	The Data Framework shall provide data visualisations.	The Data Framework shall include data visualisation with statistics and graphs about the data contained in the database which can be accessed as well as the data the user has extracted from the scenario databases. The visualisation must be easy to understand for the user. This requirement is linked to RE10 and RE11, since this requirement is about the visualisation of the data and RE10 and RE11 are about the stored information itself.
RE10	The Data Framework shall provide data statistics.	This requirement is linked to RE9 and describes the statistics about the data contained in the Data Framework the user has access to or the user has recently exported. The statistics shall include information of the scenarios, number of scenarios, parameters, parameter distributions, ODD coverage etc. The data in this requirement is the basis for the visualisation of it described in RE9.
RE11	The Data Framework shall include a data characterisation system.	This requirement describes the statistics that summarise the Data Framework's data state. The statistics shall include information of the size of the data (how many scenario databases connected, which scenarios, how

		many data per scenario type, which parameters per scenario and their distributions with mean and standard deviation, information about the scenario data contributors/uploaders). This data characterisation system shall allow the administrators and other users to get an overview of the state of the SCDBs. The linked requirement for the visualisation of these data is RE9.
RE12	Scenarios should be characterised based on their quality.	A quality metric could be based on the quality of parameter ranges and parameter combinations as well as on the transferability to test cases.
RE13	The Data Framework shall allow revisions in the data processing chain.	All components of the data processing chain in the Data Framework may have updates at a later date.
RE14	The Data Framework shall allow automated verification tests for consistency of the data format and data content.	The Data Framework must be designed to enable automated tests for data verification.
RE15	The Data Framework shall enable addition of additional scenario parameters.	Additional scenario parameters that are not included in the input data must be able to be added.
RE16	The Data Framework shall have a check for ensuring all data of the scenario elements are present during download.	This requirement is linked to RE16 and is about the existence of a check for the existence of all relevant information per scenario element (e.g., junction scenario without any road network information)
RE17	The date framework shall facilitate the retrieval of various metrics of a scenario (category) if available in scenario database across different data sources within a specific ODD.	The underlying metrics of each scenario is likely to be different. The data framework needs to ease the extraction of diverse measurements related to a given situation (category), provided they are accessible in the scenario database from different data origins within a defined ODD.
RE18	The Data Framework shall provide an option for users to store their query and query results.	This would allow the user to track and trace their queries and results. This might be useful, e.g., to reproduce results or artefacts.
RE19	The Data Framework may contain scenarios for testing connectivity and cyber-security aspects.	This would allow the user to track and trace scenarios with connectivity and cyber-security aspects.
RE20	The Data Framework shall facilitate the creation of the copies of scenarios with adjusted parameters in local platforms.	It shall be possible for the users to copy the original scenario with adjusted parameters and store it as new entry in local platform.

CL5 – Version Control Requirements:

Table 6: Version Control Requirements

Nr.	Requirements	Description
RE21	The Data Framework shall track with which version data of the Data Framework and related toolchain was created/modified.	If any of the components of the Data Framework have been changed the information on the exact version that has been used on the processed data must be available.
RE22	The Data Framework shall track the version of the scenario database.	This means if the scenario database undergoes changes to their structure, content, methodology, queries, them the data framework can track the version of the scenario database to help ensure combability and accurate data retrievable.
RE23	The Data Framework shall assign the corresponding revision version for the copied scenario.	This will facilitate the process of keeping the track of the version history of scenarios.

CL6 – File Attachment Requirements:

Table 7: File Attachment Requirements

Nr.	Requirements	Description
RE24	The Data Framework shall allow different types of files attachments.	The files are supported files and don't need to be processed such as have pictures, videos and other file formats.

CL7 – Structure of Scenario Representation Requirements:

Table 8: Structure of Scenario Representation Requirements

Nr.	Requirements	Description
RE25	The data framework should be presented the scenarios to the user by type, application and certain characteristics.	This could e.g., be a hierarchical structure, or tagging. Also, A group of concrete scenarios that share certain characteristics may be described using a logical, abstract, or functional scenario. Similarly, a group of logical scenarios that share certain characteristics may be described using an abstract or functional scenario.

CL8 – Standards Alignment Requirements:

Table 9: Standards Alignment Requirements

Nr.	Requirements	Description
RE26	The input format (from SCDBs to the data framework) shall be compatible with agreed standardised formats.	The format used to input data to the Data Framework must be compatible with currently relevant formats.
RE27	The output format of the data framework shall be compatible with other relevant formats including ASAM.	The format used to output data from the SCDB framework must be compatible with currently relevant formats (e.g., ASAM OpenDRIVE and OpenSCENARIO).

CL9 – Scenario Filtering/Searching Requirements:

Table 10: Scenario Filtering/Searching Requirements

Nr.	Requirements	Description
RE28	The Data Framework shall allow querying of scenarios via API and GUI.	One of the main purposes of the Data Framework is to enable retrieving scenarios from the different scenario databases. To enable an API or GUI to retrieve a selection of the scenarios, the data framework shall support performing queries on the different underlying scenario databases.
RE29	The Data Framework shall enable querying similar scenarios from different databases using the same query.	As per RE38, the Data Framework enables querying scenarios from different scenario databases. It would be cumbersome and impractical if the query should be different for obtaining the same (type of) scenario from different databases. To avoid this, the Data Framework shall include, if necessary, any translation of the query for individual scenario databases.
RE30	The Data Framework shall be able to query and filter based on agreed taxonomy (meta data.)	Meta data includes e.g., how the scenario was created or recorded, when it was recorded, which format it is stored in. E.g., ASAM OpenLABEL
RE31	The Data Framework shall filter scenarios based on ODD and OD.	ODD and OD includes e.g., type of traffic area, traffic conditions, environment.
RE32	The Data Framework shall filter based on omissions (NOT statement).	If the user desires a specific scenario and wants to exclude certain elements (e.g., no 3 assets or no pedestrians), the user should be able to filter it.
RE33	The result of a query with the Data Framework shall be reproducible.	For the sake of reproducibility, it is very useful if the result of a query can be reproduced. Note that this does not necessarily mean that the same query should always result in the same outcome; especially if scenarios are added to the individual scenario databases, the same query might produce more results. However,

		it should still be possible to alter the query (e.g., by adding a date range for when the scenarios have been added to the individual scenario databases) such that the same result is obtained.

CL10 – Legal and Compliance Requirements:

Table 11: Legal and Compliance Requirements

Nr.	Requirements	Description
RE34	The Data Framework shall respect the usage of raw personal information according to national and international privacy laws.	This means that privacy laws such as GDPR or laws from relevant countries should be respected and also that some scenarios might not be available in other countries, if a scenario does not match corresponding privacy laws. Furthermore, the data framework shall only accept anonymised scenario data.
RE35	Copyrights and licenses must be present and visual for the user with respect to all data sources in the pipeline.	This requirement is about the copyright and license information of the source of the exported data, the Data Framework itself and the source databases. These copyright and license information must be included in any kind of GUI and exported files.
RE36	Data framework shall comply with relevant cybersecurity best practices and standards.	This means relevant cyber security best practices (e.g. access control for authentication and authorisation) has be complied with. This will enable a robust security.
RE37	The Data Framework shall be hosted in a secure cloud environment.	Cloud hosting provides reliability and flexibility at a manageable cost.

CL11 – Scenario Metadata Association Requirements:

Table 12: Scenario Metadata Association Requirements

Nr.	Requirements	Description
RE38	The data framework shall contain scenario meta information, based on agreed taxonomy (meta data)	Meta information like data source, type of data, version, timestamp, proposed measures, manual or system driving, etc. must be included in the output data.
RE39	Copyright shall be part of the output data, based on agreed taxonomy.	This requirement is the copyright and license information of the data source that shall be contained in the output data (can be also comments in the OpenSCENARIO and OpenDRIVE files).
RE40	The data framework should include the definition of mandatory and optional parameters based on agreed taxonomy/ontology.	Including naming conventions for these parameters.

RE41	Scenario Metadata, based on agreed taxonomy, shall include: origin of the scenario data, scenario added date, whether the ADS was activated,	In order to know whether scenarios are relevant for the safety assurance of an ADS, it might be useful to know the origin of the scenario data. For example, the scenario could be made up by certain experts, the scenario could be extracted from vehicle data or observed from road-side data.

CL12 – Scenario Metrics Requirements:

Table 13: Scenario Metrics Requirements

Nr.	Requirements	Description
RE42	A scenario shall contain a definition of the output/result requirements.	Which signals are required as output (e.g. collision, lane centre offset, maximum acceleration/jerk, throughput/time.
RE43	Scenario should define which outputs are the most relevant.	This enables making a standardized overview of results from a test results database.
RE44	The metrics provided in the data framework shall be consistent with requirements from end users.	The Data Framework shall include metrics that are used for further analysis of the selected scenarios. Those metrics shall be aligned with the ones used for requirements of safety assessment, which can be defined either by end user or by bodies such as Regulators or Euro NCAP.
RE45	It shall be possible to obtain indicators on the completeness of the scenario data for a specific ODD.	For using the scenario data for the safety assurance of a system, it is important to know how complete the data is in relation to the ODD of the system.
RE46	The Data Framework shall support scenario Scoring/rating e.g., similarity, criticality, based on agreed taxonomy (meta data.)	Help filtering and scoring scenarios by identify the critical and similar scenarios.

CL13 – Scenario Definition and Format Requirements:

Table 14: Scenario Definition and Format Requirements

Nr.	Requirements	Description
RE47	The Data Framework shall be compatible with the following databases: Safety Pool™ Scenario Database, Scenius, Moove, ADScene, VVMethods, and Streetwise.	The data format from popular databases shall be readable for the Data Framework and when using scenarios, the responsibility of individual SCDBs to create mapping to the Data Framework.
RE48	The required and optional parameters/fields shall be defined.	A scenario can have specific parameters, which are not relevant for other scenarios.
RE49	The Data Framework shall support both concrete and logical scenarios.	Logical scenarios are more flexible to use, as they cover a variety of concrete scenarios, and could be used to let a

	Preference to have logical scenarios in the connected SCDBs if available.	simulation tool loop over the variations automatically.
RE50	The Data Framework shall allow definition of parameter ranges and/or parameter value sets.	To enable defining logical scenarios. Protocol scenarios can require a fixed set of parameter values. Hence a single value, a range, and a parameter set must be allowed in the database format.
RE51	The scenarios shall include the probabilities for the different parameter ranges/sets, if available	This can allow to improve the SuT for most probable cases first.
RE52	The data framework should ensure output scenarios in format that allows automated execution in simulation environments.	No manual steps should be required after getting the scenario data, up to running the simulation and obtaining simulation results.
RE53	Scenarios shall contain a human-readable description.	A brief description, so the user knows what this scenario is about.
RE54	All scenarios shall have the required fields filled with relevant data.	This ensures that the scenarios can be queried in a consistent way.
RE55	Logical scenario shall cover a variety of concrete scenarios.	This can be a range for scenario parameters or a multitude of parameter sets.
RE56	Scenario description shall be complete, such that a concrete scenario can be derived unambiguously.	Scenario definition should leave no room for interpretation.
RE57	The input format shall include dynamic objects, map data and weather information and represent changing road network information.	In order to be able to describe scenarios adequately, some information must be able to be included in any data inputted into the Data Framework.
RE58	The output file(s) shall satisfy the requests of the user for specific version and content and must be valid regarding version and content.	Abstract scenarios OpenSCENARIO >= v2.0 || logical scenarios with (offered) parameter distributions OpenSCENARIO >= v1.2 || concrete or parameterised scenarios OpenSCENARIO >= v1.0. Users can select the version of OpenScenario
RE59	The data framework shall check if the output of scenario of individual databases are in agreed format/language(s) Any scenario description language may be acceptable at SCDB level as long as the SCDB owner provides means to convert/translate or adapt to the harmonised SUNRISE scenario language(s).	This means that the output of the individual SCDBs which interact with the data framework shall be in an agreed standardised format. This format may be standard formats like ASAM OpenSCENARIO 1.X, ASAM OpenSCENARIO 2.X, BSI Flex 1889, Scenic etc.
RE60	The data framework shall enable scenarios with different formats that are: human-readable, executable and retrievable from database, if scenarios have these format.	OpenSCENARIO and OpenDRIVE are two different files/formats, so the human-readable file might also be in a different format as well.
RE61	The data framework should enable scenario definition to have optional capabilities to handle V2X, if scenarios have these format.	V2X enablers can be part of the data framework.

CL14 – Scenario Source Categorisation Requirements:

Table 15: Scenario Source Categorisation Requirements

Nr.	Requirements	Description
RE62	All data should be traceable to its origin (original input data).	Data that is processed by the Data Framework must be able to be traced to its source scenario database
RE63	The data framework shall allow feeding of data coming from various data sources.	This allows data enrichment, including also data gathered in past EU projects (e.g., L3Pilot, Hi-Drive).

CL15 – Scenario Tagging/Labelling Requirements:

Table 16: Scenario Tagging/Labelling Requirements

Nr.	Requirements	Description
RE64	The data Framework shall allow flexibility depending on end user of “results”.	E.g., reasonably foreseeable scenario is within ALKS but maybe not within Euro NCAP There may be different safety assessment needs, depending on the origin of the requirement (e.g. Regulation or Euro NCAP, or own criteria). This shall be reflected in the data framework.
RE65	The data framework shall contain a labelled scenario database. SCDBs connected to the data framework shall contain labelled scenarios	The labelled scenario database consists of scenarios with tags and labels and describes the events in a scenario, means scenario data and meta data.
RE66	The data framework may be able to auto validate scenarios based on ODD tags and behaviour tags (for scenario parameters) defined by ASAM OpenLabel tags, based on agreed ontology	When the user retrieves a scenario, the data framework should check that the tags are genuine.

CL16 –Test Results Requirements:

Table 17: Test Results Storage Requirements

Nr.	Requirements	Description
RE67	It may be possible for a user to store results of tests for the scenarios in a common format via the data framework in any of the connected SCDBs.	Test results are not stored in the Scenario data framework itself, but in a user specific database closely linked to the scenario database.
RE68	It may be possible to retrieve earlier stored test results via the data	Test results can be retrieved by searching for scenario, test date, etc.

	framework from the connected SCDBs.	
RE69	Search of test data may be linked to the scenario search which is done via the data framework	This e.g. would allow searching for scenarios that have not been tested yet.
RE70	The data framework may provide an overview of the main result parameters for retrieved test results.	When searching for test results, the data framework could e.g., show a table with a summary of the main results.

Overall, the list has a mix of requirements, ranging from technically specific to more general in nature. While some specifications are clear-cut, others serve broader purposes, such as usability enhancements or operational efficiencies.

3.5 Refined Feeback

Further feedback has been collected, highlighting key enhancements to fortify the data framework. This feedback emphasizes the importance of delineating incentives for stakeholders, clarifying their motivations for utilizing the data framework. Additionally, it underscores the necessity to define the hosting cloud methodology for the data framework. Furthermore, there is a need to specify the utility of the framework for scientific stakeholders, elucidating whether access aligns with industrial usage or if it's customized for scientific purposes.

4 CONCLUSIONS

In this deliverable, detailed and comprehensive requirements are defined for the SUNRISE data framework and federated scenario databases, by including the interfacing between the data framework and the scenario databases that connect to it, searching for scenarios, presenting scenarios to the end user, and the storing/retrieving of test results. Additionally, a clustering of the requirements was made according to their topic or theme. The methodology used, led to the production of 71 distinct requirements, distributed in 16 clusters (see section 3). Each of the aforementioned distinct requirements is uniquely tagged to be referred by other WPs within the SUNRISE project, such as WP7 on use cases and framework demonstration instances development. The gathered requirements for the data framework cover the standards and governance, and a set of rules of scenario definition and format. This will be essential for the necessary provisions for seamlessly accommodating future scenarios and their parameters as part of the data framework and constituting the scenario databases.

The requirements have been gathered through collaboration among partners from diverse backgrounds, encompassing academia, research, and industry. This effort yielded a comprehensive set of requirements for the utilization of scenario data in ensuring the safety of Connected and Automated Mobility (CCAM) systems. Moreover, during the survey efforts involving stakeholders in the ecosystem, more than 50 organizations were contacted, ensuring a thorough refinement of these requirements (See Annex 1).

The present work can also be utilised mainly in Task 5.2 and Task 5.3, which are related to data framework harmonisation, quality metrics development, interface synchronisation, and architecture development. Furthermore, this effort will aid in the completion of Task 6.1 and Task 6.2 by establishing guidelines for the format of scenario inputs and outputs within the data framework.

5 REFERENCES

1. A. Thorsén, A. Farooqui, M.Skoglund, P. Stålberg, D. Becker, E. Arnoux, O. Op den Camp, P. Irvine, C. Berger, T. Bouraffa, C. Kaya, M. Muro, G. Ben Nejma, O. Bartels, F. Warg, A. Bolovinou, I. Panangiotopoulos, “Sunrise D3.1 Report on baseline analysis of existing Methodology”, 2023

2. ISO (2022) ISO 34501:2022 - Road vehicles -- Test scenarios for automated driving systems – Vocabulary

3. “PEGASUS project family”, https://pegasus-family.de/

4. “SAKURA project”, https://www.sakura-prj.go.jp/

5. “Streetwise”, https://www.tno.nl/en/digital/smart-traffic-transport/smart-vehicles/integrated-vehicle-safety-smart-vehicles/streetwise/

6. “Safety Pool Database.” https://www.safetypool.ai/ (accessed Aug. 29, 2023).

7. “ADScene”, https://dsc2021.org/adscene-towards-an-industrial-scenarios-plateform-for-driving-assistance-systems-design-validation/

8. “ArchitectECA2030”, https://architect-eca2030.eu/

9. Elrofai, H., Paardekooper, J.-P., de Gelder, E., Kalisvaart, S., and Op den Camp, O. (2018). Scenario-based safety validation of connected and automated driving. Netherlands Organisation for Applied Scientific Research, TNO.

10. 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 2021 IEEE ITSC, 2021, pp. 312–319.

11. X. Zhang, S. Khastgir, and P. Jennings, “Scenario Description Language for Automated Driving Systems: A Two Level Abstraction Approach,” in 2020 IEEE SMC, 2020, pp. 973–980.

12. A. Anastasio Bruto da Costa, P. Irvine, X. Zhang, S. Khastgir, and P. Jennings, “Translating Automated Vehicle Test Scenario Specifications Between Scenario Languages: Learnings and Challenges,” in Proceedings of the Driving Simulation Conference 2022 Europe VR, 2022, pp. 65–72.

13. P. Irvine, P. Baker, Y. K. Mo, A. B. Da Costa, X. Zhang, S. Khastgir, and P. Jennings, “Vehicle-to-Everything (V2X) in Scenarios: Extending Scenario Description Language for Connected Vehicle Scenario Descriptions,” in 2022 IEEE IV, 2022, pp. 548–555.

14. BSI, “BSI Flex 1889 - Natural language description for abstract scenarios for automated driving systems - Specification,” 2022.

15. X. Zhang, S. Khastgir, and P. Jennings, “An ODD-Based Scalable Assurance Framework for Automated Driving Systems,” in SAE WCX. SAE International, 2023.

16. ISO, “ISO 34503 - Road Vehicles — Test scenarios for automated driving systems — Specification for operational design domain,” 2023.

17. “ASAM OpenODD”, [Online]. Available: https://www.asam.net/standards/detail/openodd/

18. BSI, “PAS 1883 Operational Design Domain (ODD) taxonomy for an Automated Driving System (ADS) Specification.” 2020.

19. A. Wallace, S. Khastgir, X. Zhang, S. Brewerton, B. Anctil, P. Burns, D. Charlebois, P. Jennings “Validating Simulation Environments for Automated Driving Systems Using 3D Object Comparison Metric,” in _2022 IEEE Intelligent Vehicles Symposium (IV)_, 2022, pp. 860–866. doi: 10.1109/IV51971.2022.9827354.

20. UNECE, Validation Method for Automated Driving (VMAD).

21. ASAM e.V., “OpenLABEL V1.0.0,” 2021. https://www.asam.net/standards/detail/openlabel/ (accessed Mar. 29, 2023).

ANNEX 1: GATHERING METHODOLOGIES

The methodologies used to gather the requirements used by the different SUNRISE participants are listed below:

BASt: The majority of requirements were extracted from the results of an internal project that developed an architecture and role model for the operation of a successful scenario database in Germany. Relevant requirements were extracted and translated to requirements for the data framework of SUNRISE. Further ones were developed on the existing ones to improve the analysis capabilities for different scenarios. As such a data framework is vital for regulators to develop regulations, many requirements concern the success of the data framework.

ICCS: The approach to extract requirements stood thematically around connectivity and cyber-security issues, since they are both main aspects to enrich a SCDB with essential information. Procedurally, an internal consultation round concluded to several statements around the above issues, which consequently were brought to the task level and aligned accordingly, taking their final structure.

IDIADA: The approach followed to extract requirements has been based on the extensive experience of IDIADAs' team on dealing with safety related matters from a technical service perspective. Internal discussions within the team have been the main tool to identify the most critical requirements for the SUNRISE's federated database and its utilisation. This database will be crucial for type-approval related activities; therefore, the success and adoption of the database have been a major point while deriving the requirements.

Ika: The requirements were gathered based on the extensive experience that ika has gained through the development and maintenance of existing scenario databases in projects such as Pegasus and VVM. To accomplish this, expert knowledge was utilised to identify the most pertinent requirements for these databases and their utilisation within the comprehensive federated scenario database layer envisioned for the SUNRISE project.

Siemens: The requirements have been gathered by brainstorming within our team. Some structure to the brainstorm session was given by grouping requirements in 3 levels (Data framework, scenario database format, and scenario database contents). Finally, the obtained set of requirements was reviewed by the team.

Toyota Motor Europe: The approach followed has been based on the experience that our team has in dealing with data for safety assessment, trying to clarify requirements at different levels, starting from type of data used and formatting, up to the need to understand what the data framework purpose is for as well as the SCDB content. The requirements may differ depending on the usage purpose, so TME referred to experience from the SAKURA project which also developed an SCDB for safety assessment in Japan. This project included aspects such as the need to clarify users, purpose, requirements of the SCDB content as well as input data used.

TNO: TNO obtained requirements for the data framework using an internal discussion with experts working on the data-driven scenario database StreetWise [9]. These experts have extensive contacts with partners that are using the StreetWise methods to extract real-world scenarios from data and to collect those scenarios into a database that is used for the verification and validation of ADSs. In addition to this, stakeholders of the StreetWise SCDB, such as road authorities, consumer organisations, and OEMs have been consulted for acquiring additional requirements for the data framework.

University of Warwick (UoW): UoW has complied the requirements based on our research and standardisation activities in Safety assurance [10], scenario-based testing [11][12][13][14], ODD [15][16][17][18], virtual test environment validation [19][20], databased development [21]. In addition, our network with various stakeholder and expert groups through our R&D activities and Safety Pool™ [6] initiative enabled us to form a broad and comprehensive view point, they include regulators, research institute, Tier-1 supplier, OEMs, system developers, test engineers. The safety assurance process contains scenario, execution environment, and analysis at a high level, underpinned by ODD and behaviour. As part of survey activities of the various stakeholders in the ecosystem, over 50 organisations were contacted through our engagement via a Safety Pool™ Scenario Database booth at the Driving Simulation Conference 2022 (Sep 2022), ASAM International Conference 2022 (Nov 2022), and Autonomous Vehicle Test & Development Expo Stuttgart 2023 (June 2023).

Vicomtech: The proposed requirements have been gathered by a team of researchers with experience in scenario-based testing, aiming to cover the needs of research and industrial partners on the utilisation of the SUNRISE data framework. Where possible, requirements have been detailed so as to consider specific technical challenges, such as data formats. All requirements have been formulated as natural language sentences that can be interpreted as imperative requirements the platform shall fulfil. Such format simplifies interpretation and minimises ambiguities.

External stakeholder engagement: The engagement with external stakeholders has been conducted through various processes, emphasizing collaboration and feedback collection. These processes include standardization efforts, collaborative projects, dissemination activities, and participation in national/international events such as ADAS & Autonomous Vehicle Technology Expo in Stuttgart, Germany; the Driving Simulation and Virtual Reality Conference in Antibes, France; and the IPG Automotive Open House in Coventry, United Kingdom. While these stakeholders are not formal project partners, their input is considered crucial. Feedback has been gathered through informal channels, primarily verbal exchanges, and subsequently integrated into the requirement table. Although maintaining traceability with such informal feedback channels can be challenging, the insights obtained are deemed valuable for the project's progress.