PD ISO/TR 4804:2020

BSI Standards Publication

Road vehicles — Safety and cybersecurity for automated driving systems — Design, verification and validation

Documents
Document Sharing Hub

Full document details: PD ISO/TR 4804:2020

National foreword

This Published Document is the UK implementation of ISO/TR 4804:2020.

The UK participation in its preparation was entrusted to Technical Committee AUE/32, Electrical and electronic components and general system aspects (Road vehicles).

A list of organizations represented on this committee can be obtained on request to its committee manager.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2021
Published by BSI Standards Limited 2021

ISBN 978 0 539 12636 5

ICS 43.020

Compliance with a British Standard cannot confer immunity from legal obligations.

This Published Document was published under the authority of the Standards Policy and Strategy Committee on 28 February 2021.

Amendments/corrigenda issued since publication

Date	Text affected

Documents
Document Sharing Hub

Full document details: PD ISO/TR 4804:2020

TECHNICAL REPORT

ISO/TR 4804

First edition
2020-12

Road vehicles — Safety and cybersecurity for automated driving systems — Design, verification and validation

Véhicules routiers — Sécurité et cybersécurité pour les systèmes de conduite automatisée — Conception, vérification et validation

Reference number
ISO/TR 4804:2020(E)

Documents
Document Sharing Hub

© ISO 2020

Full document details: PD ISO/TR 4804:2020

COPYRIGHT PROTECTED DOCUMENT

© ISO 2020

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub

Full document details: PD ISO/TR 4804:2020

Contents

Page

Foreword ..... **v**
Introduction ..... **vi**
1 Scope **1**
2 Normative references **1**
3 Terms and definitions **1**
4 General approach and overview **11**
4.1 Introduction and motivation 11
4.2 Overview of this document 11
4.3 Structure and development examples used in this document 12
4.4 Safety vision 13
4.4.1 Background 13
4.4.2 Positive risk balance and avoidance of unreasonable risk 14
4.4.3 Principles of safety and cybersecurity for automated driving 14
5 Systematically developing dependability to support safety by design **17**
5.1 General 17
5.2 Deriving capabilities of automated driving from dependability domains 18
5.2.1 Applying the related safety standards 18
5.2.2 ISO/PAS 21448 - Safety of the intended functionality 19
5.2.3 ISO 26262 series - Functional safety 19
5.2.4 ISO/SAE 21434 - Automotive cybersecurity 20
5.2.5 Capabilities of automated driving 21
5.2.6 Minimal risk conditions and minimal risk manoeuvres 25
5.3 Elements for implementing the capabilities 27
5.3.1 Implementing the capabilities 27
5.3.2 Elements 33
5.3.3 Generic logical architecture 45
6 Verification and validation **48**
6.1 General 48
6.2 The scope and main steps of verification and validation for automated driving systems 49
6.3 Key challenges for verification and validation of SAE L3 and SAE L4 automated driving systems 50
6.3.1 Challenge 1: Statistical demonstration of avoidance of unreasonable risk and a positive risk balance without driver interaction 51
6.3.2 Challenge 2: System safety with driver interaction (especially in takeover manoeuvres) 51
6.3.3 Challenge 3: Consideration of scenarios currently not known 51
6.3.4 Challenge 4: Validation of various system configurations and variants 51
6.3.5 Challenge 5: Validation of (sub)systems that are based on machine learning 51
6.4 Verification and validation approach for automated driving systems 51
6.4.1 Defining test goals and objectives (why and how well) 52
6.4.2 Test design techniques (how) 52
6.4.3 Test platforms (where) 53
6.4.4 Test strategies in response to the key challenges 53
6.5 Quantity and quality of testing 57
6.5.1 Equivalence classes and scenario-based testing 58
6.6 Simulation 58
6.6.1 Types of simulation 60
6.6.2 Simulation scenario generation 61
6.6.3 Validating simulation 61
6.6.4 Further applications of simulation 62
6.7 Verification and validation of elements 62
6.7.1 A-priori information and perception (map) 63

AVA Documents
Document Sharing Hub

© ISO 2020 – All rights reserved

Full document details: PD ISO/TR 4804:2020

6.7.2 Localization (including GNSS) 63
6.7.3 Environment perception sensors, V2X and sensor fusion 64
6.7.4 Interpretation and prediction, drive planning and traffic rules 64
6.7.5 Motion control 65
6.7.6 Monitor, ADS mode manager (including the vehicle state) 65
6.7.7 Human machine interaction and user state monitor 65
6.8 Field operation (monitoring, configuration, updates) 65
6.8.1 Testing traceability 65
6.8.2 Robust configuration and change management process 66
6.8.3 Regression prevention 67
6.8.4 Cybersecurity monitoring and updates 67
6.8.5 Continuous monitoring and corrective enforcement 67

Annex A (informative) **Development examples** **69**

Annex B (informative) **Using deep neural networks to implement safety-related elements for automated driving systems** **80**

Annex C (informative) **Principles of safety and cybersecurity for automated driving** **92**

Annex D (informative) **List of proposed standards** **95**

Bibliography **107**

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 22, *Road Vehicles*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

Introduction

Automated driving is one of the key modern technologies. In addition to offering broader access to mobility, it may also help to reduce the number of road traffic related accidents and crashes. When doing so, the safe operation of automated driving vehicles is one of the most important factors. Designed to supplement existing standards and publications on various aspects of safety, this document presents a more technical overview of the recommendations, guidance and methods to achieve a positive risk balance and to avoid unreasonable risk and cybersecurity related threats, emphasizing the importance of safety by design. This document closes the loop to provide a discussion with recommendations and methods on the verification and validation of automated driving systems.

Set forth are a proposed framework and guidelines focused on the safety and cybersecurity during the development, verification, validation, production and operation of automated driving systems for all stakeholders in the automotive and mobility world – from technology start-ups through to established OEMs and the tiered suppliers of key technologies.

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

Road vehicles — Safety and cybersecurity for automated driving systems — Design, verification and validation

1 Scope

This document describes steps for developing and validating automated driving systems based on basic safety principles derived from worldwide applicable publications. It considers safety- and cybersecurity-by-design, as well as verification and validation methods for automated driving systems focused on vehicles with level 3 and level 4 features according to SAE J3016:2018. In addition, it outlines cybersecurity considerations intersecting with objectives for safety of automated driving systems.

2 Normative references

There are no normative references in this document

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1
automated driving system
ADS
set of *elements* (3.14) that offer a specific conditional or higher automated driving *use case* (3.63) in or for a specific *ODD* (3.37)

3.2
automated vehicle
AV
vehicle equipped with at least one conditional (SAE level 3) or higher (SAE level 4/level 5) *automated driving system* (3.1)

3.3
availability
capability (3.4) of a product to provide a stated function if demanded, under given conditions over its defined lifetime

Note 1 to entry: In the context of this document the product is the *automated driving system* (3.1).

Note 2 to entry: In the context of this document "availability" is defined solely referring to the automated driving system aspects and does not include human factor aspects.

[SOURCE: ISO 26262-1:2018, 3.7]

3.4
capability
ability of a product to deliver a function, feature or service

Note 1 to entry: In the context of this document the product is the *automated driving system* (3.1).

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

3.5
conventional driver
driver (3.11) who manually exercises in-vehicle braking, accelerating, steering and transmission gear selection input devices in order to operate the vehicle

[SOURCE: SAE J3016:2018, 3.29.1.1]

3.6
corner case
scenario (3.53) in which two or more parameter values are each within the *capabilities* (3.4) of the system, but together constitute a rare condition that challenges its capabilities

Note 1 to entry: In the context of this document the system is the *automated driving system* (3.1).

[SOURCE: ISO/PAS 21448:2019, Table 11]

3.7
crash
undesirable, unplanned event that leads to an unrecoverable loss due to unfavourable external conditions (e.g. human error), typically involving material damage, financial loss or human injuries and/or fatalities

3.8
cybersecurity
condition in which assets are sufficiently protected against threat *scenarios* (3.53) to electrical or electronic components of road vehicles and their functions

[SOURCE: ISO/SAE 21434]

3.9
degradation
state or transition to a state of the *item* (3.26) or *element* (3.14) with reduced functionality, performance, or both

Note 1 to entry: In the context of this document the item is the *automated driving system* (3.1).

[SOURCE: ISO 26262-1:2018, 3.28, modified — Note 1 to entry added.]

3.10
dependability
ability of a system to provide a service or function regarding the attributes of *reliability* (3.44), *availability* (3.3), maintainability, *safety* (3.51) and security (RAMSS)

Note 1 to entry: In the context of this document the system is the *automated driving system* (3.1).

3.11
driver
user (3.64) who performs in real-time part or all of the *DDT* (3.13) and/or DDT fallback for a particular vehicle

[SOURCE: SAE J3016:2018, 3.29.1, modified — The word "human" was removed from the term and the note was deleted.]

3.12
driver in the loop
DiL
execution of the target software on prototype or target hardware in the target vehicle or a mock-up, in which the environment is modified with virtual stimuli, and the driver's reaction influences the vehicle's behaviour

EXAMPLE Driving simulator or vehicle in the loop (ViL) (augmented reality for safety-related manoeuvres in real vehicles).

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

3.13
dynamic driving task
DDT
all of the real-time operational and tactical functions required to operate a vehicle in on-road traffic

Note 1 to entry: This excludes the strategic functions such as trip scheduling and selecting destinations and waypoints, and includes without limitation:

— lateral vehicle motion control via steering (operational);

— longitudinal vehicle motion control via acceleration and deceleration (operational);

— monitoring the driving environment via object and event detection, recognition, classification, and response preparation (operational and tactical);

— object and event response execution (operational and tactical);

— manoeuvre planning (tactical); and

— enhancing conspicuity via lighting, signalling or gesturing, etc. (tactical).

[SOURCE: SAE J3016:2018, 3.13, modified — Note 1 to entry was previously part of the definition, the notes, figure and additional information were removed.]

3.14
element
at least first-level decomposition of *capabilities* (3.4) to a logical system architecture

Note 1 to entry: One or more elements realize one or more capabilities.

3.15
equivalence class
class being identified based on the division of inputs and outputs, such that a representative test value can be selected for each class

Note 1 to entry: See ISO 26262-6:2018, Table 8.

3.16
fail-degraded
property of the *item* (3.26) to operate with reduced functionality in the presence of a *fault* (3.20)

Note 1 to entry: This property can be realized as fail-degraded *capability* (3.4) of fail-degraded mode.

Note 2 to entry: In the context of this document the item is the *automated driving system* (3.1).

Note 3 to entry: This means that the item is fault-tolerant for a subset of its intended functionality.

Note 4 to entry: The absence of *unreasonable risk* (3.62) can require the duration of the presence of the fault to be time limited and/or system maintenance in a limited time frame.

Note 5 to entry: The absence of unreasonable risk in the presence of the fault can require limitations of the item behaviour.

3.17
fail-operational
property of the *item* (3.26) to maintain its full intended functionality in the presence of a *fault* (3.20)

Note 1 to entry: In the context of this document the item is the *automated driving system* (see 3.1).

Note 2 to entry: This means that the item is fault-tolerant for its intended functionality.

Note 3 to entry: The absence of *unreasonable risk* (3.62) can require the duration of the presence of the fault to be time limited and/or system maintenance in a limited time frame.

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

3.18
fail-safe
property of an *automated driving system* (3.1) to achieve a *minimal risk condition* (3.29) and to achieve a *safe state* (3.50) in the event of a *failure* (3.19)

Note 1 to entry: A fail-safe condition is to be reached for example, by means of: demanding the vehicle control to driver/vehicle *operator* (3.39) and/or switching off the automated driving function.

3.19
failure
termination of an intended behaviour of an *element* (3.14) or the *automated driving systems* (3.1) due to a *fault* (3.20) manifestation

[SOURCE: ISO 26262-1:2018, 3.50, modified — The term "automated driving system" replaces "item" and Note 1 to entry is not included here.]

3.20
fault
abnormal condition that can cause an *element* (3.14) or the *automated driving system* (3.1) to fail

[SOURCE: ISO 26262-1:2018, 3.54, modified — The term "automated driving system" replaces "item" and Notes to entry are not included here.]

3.21
field operational testing
FOT
use of large-scale testing programs aimed at generating a comprehensive assessment of the efficiency, quality, robustness and acceptance of transport solutions

3.22
high definition map
HD map
maps with high level precision mostly used in the context of *automated driving system* (3.1) to give the vehicle precise information about the road environment

3.23
hardware in the closed loop
HiL
execution of target software on target hardware, whereby the hardware outputs influence the hardware inputs

Note 1 to entry: HiL executes the target software in real time.

EXAMPLE AUTOSAR stack on radar with no frontend.

3.24
hardware open loop
HoL
execution of target software on target hardware, whereby the hardware outputs do not influence the hardware inputs

EXAMPLE Monitor hardware testbench.

3.25
human-machine interaction
interdisciplinary interaction between a human and an *automated vehicle* (3.2), considering the human-machine interface (HMI) with the aim to develop a user interface that satisfies requirements regarding mental, cognitive and manual abilities of the *user* (3.64)

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

3.26
item
system or combination of systems, that implements a function or part of a function at the vehicle level

[SOURCE: ISO 26262-1:2018, 3.84, modified — The phrase "to which ISO 26262 is applied" and the Note 1 to entry were deleted.]

3.27
lagging measure
metrics that are assessed after deployment of an *automated driving system* (3.1) and provide confirmation that the *positive risk balance* (3.42) as well as the conformance with the safety-by-design techniques have been achieved

EXAMPLE Statistics for *crashes* (3.7) or other *safety* (3.51) events.

Note 1 to entry: See Reference [1].

3.28
leading measure
metrics that are derived from data that is assessed prior to deployment of an *automated driving system* (3.1) indicating that the automated driving system conforms with safety-by-design techniques to achieve a *positive risk balance* (3.42) and avoidance of *unreasonable risk* (3.62)

EXAMPLE A design *verification* (3.67) that the HMI guidelines were incorporated into the vehicle's design.

Note 1 to entry: See Reference [1].

3.29
minimal risk condition
MRC
condition to which a *user* (3.64) or an *automated driving system* (3.1) may bring a vehicle after performing the *minimal risk manoeuvre* (3.30) in order to reduce the risk of a *crash* (3.7) when a given trip cannot be completed

Note 1 to entry: The minimal risk condition integrates the meaning of avoidance of *unreasonable risk* (3.62), according to the ISO 26262:2018 series. They can be combined but they never exclude one each other.

[SOURCE: SAE J3016:2018, 3.17, modified — The term "minimal risk manoeuvre" replaces "DDT fallback", the notes and examples were deleted and the Note 1 to entry was added.]

3.30
minimal risk manoeuvre
MRM
automated driving system's (3.1) *capability* (3.4) of transitioning the vehicle between nominal and *minimal risk conditions* (3.29)

3.31
operating mode awareness
driver's (3.11) *capability* (3.4) to identify the current automation mode and his/her driving responsibility

3.32
naturalistic driving study
NDS
driving study where research subjects are recruited to drive on public roads (not in a simulator or on a test track), where there is no in-vehicle experimenter or confederate vehicles, and where driving conditions are not experimentally controlled or manipulated

Note 1 to entry: Subjects are not instructed to drive differently than they normally would and the data collection instrumentation is unobtrusive.

Note 2 to entry: Typically, these studies last a minimum of several weeks for each subject and can go much longer.

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

Note 3 to entry: An approach during which the driver becomes unaware of observation as data is collected as discreetly as possible. This data is then used to examine the relationship between the driver, vehicle and/or environment.

3.33
nominal performance
performance of the system free from *fault* (3.20) and that meets its defined performance criteria

3.34
non-vulnerable road user
protected *road users* (3.46) such as *users* (3.64) in other vehicles, trucks, construction and agricultural machines

3.35
object under test
OuT
item (3.26) or *element* (3.14) to be tested as planned and specified

Note 1 to entry: Similar usage as ISO 16750 for device under test.

3.36
open road testing
execution of target software on target hardware in the target vehicle with a *driver* (3.11), whereby the driving environment is real, road infrastructures are public and can be only partially controlled

EXAMPLE *Field operational testing* (3.21) or *naturalistic driving studies* (3.32), testing in the development vehicles.

3.37
operational design domain
ODD
operating conditions under which a given *automated driving system* (3.1) or feature thereof is specifically designed to function, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics

Note 1 to entry: These limitations, as from constraints as specified in operating conditions, reflect the technological *capability* (3.4) of the automated driving system.

[SOURCE: SAE J3016:2018, 3.22, modified — Note 1 to entry is added, and the note and examples are not included here.]

3.38
ODD functional adaptation
operational design domain functional adaptation
property of a system to operate safely with reduced performance in the case of detected functional insufficiencies inside the *ODD* (3.37)

EXAMPLE Speed adaptation because of low fuel level, or dense fog, or sensor performance insufficiencies.

3.39
operator
designated person, appropriately trained and authorized, to operate the vehicle

3.40
other road user
vulnerable road users (3.68) and *non-vulnerable road users* (3.34) with no role in the ego *automated vehicle* (3.2)

3.41
passenger
user (3.64) in a vehicle who has no role in the operation of that vehicle

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

3.42
positive risk balance
benefit of sufficiently mitigating residual risk of traffic participation due to *automated vehicles* (3.2)

Note 1 to entry: This includes the expectation that automated vehicles cause less *crashes* (3.7) on average compared to those made by drivers.

Note 2 to entry: Positive risk balance is one of the concepts that can be considered when defining the acceptance criteria of ISO/PAS 21448:2019.

3.43
proving ground testing
execution of target software on target hardware in the target vehicle in a realistic but controlled and private driving environment

Note 1 to entry: The driver can be real or a robot.

EXAMPLE Emergency braking assistant tests on soft crash target.

3.44
reliability
ability of a system to continuously provide correct service

3.45
reprocessing
replay of time stamped, recorded data to provide input for the *object under test* (3.35)

Note 1 to entry: The time stamp needs a sufficient time accuracy.

3.46
road user
anyone who uses a road including sidewalk and other adjacent spaces

Note 1 to entry: Relationship between the human related terms are shown in Figure 1.

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

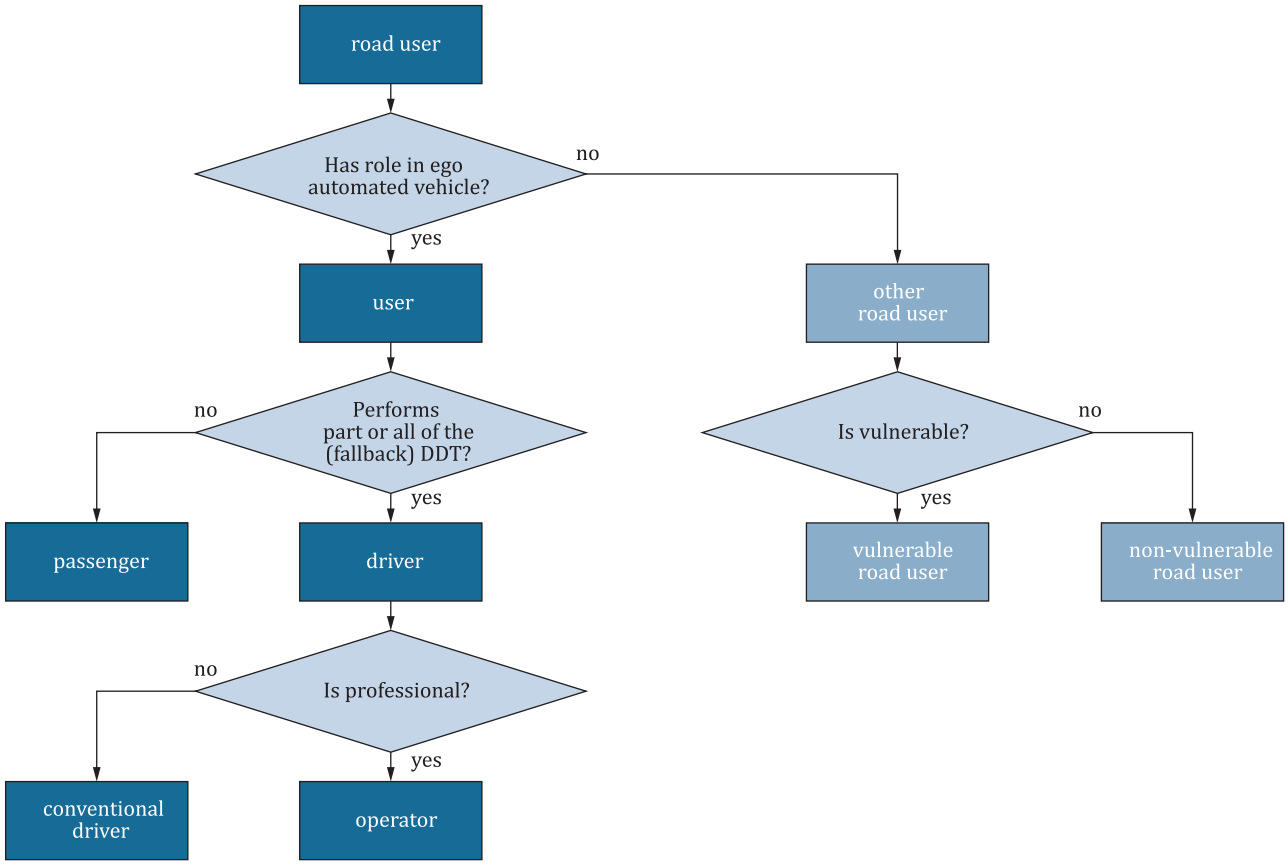

Figure 1 — Decision tree for terms related to automated driving systems

3.47
remote driver
driver (3.11) who is not seated in a position to manually exercise in-vehicle braking, accelerating, steering and transmission gear selection input devices (if any) but is able to operate the vehicle

[SOURCE: SAE J3016:2018, 3.29.1.2, modified — The notes and examples were removed.]

3.48
remote operator
operator (3.39) who is not seated in a position to manually exercise in-vehicle braking, accelerating, steering and transmission gear selection input devices (if any) but is able to operate the vehicle with or without direct vision

3.49
SAE levels of driving automation
levels of driving automation as defined in SAE J3016

Note 1 to entry: Six levels of driving automation, ranging from no driving automation (level 0) to full driving automation (level 5), in the context of motor vehicles and their operation on roadways. See Table 1.

Table 1 — Level of automation SAE level

Level	Level of automation
Driver performs part or all of the DDT	
0	No driving automation
1	Driver assistance
2	Partial driving automation

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

Table 1 *(continued)*

Level	Level of automation
ADS performs the entire DDT (while engaged)	
3	Conditional driving automation
4	High driving automation
5	Full driving automation

3.50
safe state
operating mode that is reasonably safe

Note 1 to entry: The safe state is the state in which both *fail-safe* (3.18) and *fail-degraded* (3.16) systems will provide a solution (technically provided by an alternative functionality) to avoid risk, in an acceptable criterion, to any *road user* (3.46).

3.51
safety
absence of *unreasonable risk* (3.62)

[SOURCE: ISO 26262-1:2018, 3.132]

3.52
safety of the intended functionality
SOTIF
absence of *unreasonable risk* (3.62) due to hazards resulting from functional insufficiencies of the intended functionality or by reasonably foreseeable misuse by *road users* (3.46)

[SOURCE: ISO/PAS 21448:2019, modified — The term "road users" replaces “persons”.]

3.53
scenario
description of the consecutive time series of activities integrating the subject vehicle, all its external environment and their interactions in the process of performing a certain driving task

3.54
scene
snapshot of the *scenario* (3.53) at a given point of time

Note 1 to entry: Only a simulated scene can be all-embracing (i.e. objective, otherwise known as ground truth), whereby a real-world scene is incomplete, afflicted with *faults* (3.20) and uncertainties, and observed from a subjective perspective.

3.55
scenery
part of the environment that remains unchanged

3.56
simulation
approximated imitation of selected behavioural characteristics of one physical or abstract system by a static or dynamic model

Note 1 to entry: The simulation represents the behaviour over time in which the system or parts of it are replaced by the model. It includes *SiL* (3.58), *SoL* (3.57), *HiL* (3.23), *HoL* (3.24) and *DiL* (3.12).

3.57
software reprocessing open loop
SoL
execution of target software on hardware, whereby the software decisions have no influence on the stimulus

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

3.58
software in the closed loop
SiL
execution of partial target software on hardware, whereby the software decisions influence the virtually generated stimulus

EXAMPLE MATLAB Simulink model, AUTOSAR Stack, C++ dynamic link library.

3.59
system limit
limit of the operation as stated in the *ODD* (3.37) for the specific system of interests

3.60
takeover
transfer of the driving task between the *automated driving system* (3.1) and the *driver* (3.11)

Note 1 to entry: The term handover is not used in this document. In this context the terms handover and takeover are regarded as synonyms.

3.61
tele operator
remote operator (3.48) without direct vision, but with tele-transmitted information (e.g. by cameras)

3.62
unreasonable risk
risk judged to be unacceptable in a certain context according to valid societal moral concepts

[SOURCE: ISO 26262-1:2018, 3.176]

3.63
use case
specification of a generalized field of application, possibly entailing information of the system on one or several *scenarios* (3.53), the functional range, the desired behaviour and the *system limits* (3.59)

Note 1 to entry: The use case description typically does not include a detailed list of all relevant scenarios for this use case. Instead a more abstract description of these scenarios is used.

3.64
user
general term referencing the human role in driving automation

[SOURCE: SAE J3016:2018, 3.29, modified — The word "human" was removed from the term and the notes were deleted.]

3.65
vehicle-to-everything
V2X
technology that allows a vehicle to exchange additional information with infrastructure, other vehicles and *other road users* (3.40)

Note 1 to entry: V2X can provide a growing number of helpful information such as parking space availability, upcoming road hazards and map updates, or support tele-operation of the *automated vehicle* (3.2) in relevant *scenarios* (3.53).

3.66
validation
confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled

[SOURCE: ISO/IEC/IEEE 15288:2015, 4.1.53, modified — Note 1 to entry was deleted.]

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

3.67
verification
confirmation, through the provision of objective evidence, that specified requirements have been fulfilled

Note 1 to entry: Typically used to obtain fast feedback during development.

[SOURCE: ISO/IEC/IEEE 15288:2015, 4.1.54, modified — Note 1 to entry was replaced.]

3.68
vulnerable road user
VRU
non-protected *road user* (3.46) such as motorcyclists, cyclists, pedestrians and persons with disabilities or reduced mobility and orientation

4 General approach and overview

4.1 Introduction and motivation

Automated driving is one of the key modern technologies. In addition to offering broader access to mobility, it may also help to reduce the number of driving-related crashes. When doing so, the safe operation of automated vehicles is one of the most important factors. Designed to supplement existing publications and standards (see Annex D) on various aspects related to safety and cybersecurity of automated driving, this document presents a more technical overview of the specification for the development to avoid safety-related hazards, emphasizing the importance of safety by design. Furthermore, this document aims to provide a sound discussion of the verification and validation of such systems.

This document is intended to contribute to current activities working towards the industry-wide standardization of design considerations and their verification for automated driving systems. This effort will also contribute toward a deeper understanding by developing a framework or guideline for the safety of automated driving systems for all stakeholders in the automotive and mobility world – from technology start-ups to established OEMs and the tiered suppliers of key technologies.

4.2 Overview of this document

The goal of this document is to provide an overview of and guidance about the general steps for developing, verifying and validating automated driving system safety. The starting point for this is the extraction of principles taken from different regulatory publications (various legal frameworks from around the world, ethics reports, etc.). These principles are the foundation of this document, forming the basis from which the safety by design methods and verification and validation strategies are derived. The constant focus hereby is on the development that is required in addition to existing SAE L1 and L2 driver assistance systems. It is important to consider cybersecurity in conjunction with safety, whereby cybersecurity evaluation will necessitate the use of additional analysis and technical mechanisms which in turn impacts safety. Thus, safety and cybersecurity need work together, and 5.2.4 explores this concept in detail.

This document further aims to develop guidance to tackle the risks introduced by automated driving system. It reflects the state-of-the-art automated driving technology and capabilities at the time of publication. As such, it is not a complete work and its contents will be revisited and revised as advances are made in the areas of social acceptance, technology, legal and legislation.

Devising an explicit technical solution or minimum or maximum standard is not included in the scope of this document, as several possibilities exist regarding the definition of the automated driving system, its operational design domain and technical advancements, and unique technical design elements that vary among manufacturers and that affect system capabilities, etc. Due to its focus, this document does not address topics such as abuse, data privacy and advanced driver assistance systems. Finally, non-safety-relevant elements that are normally part of a customer function, such as a comfortable driving

© ISO 2020 – All rights reserved

AVA Documents Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

strategy or the fastest navigation from point to point, are also not explicitly considered in the scope of this document.

4.3 Structure and development examples used in this document

This document is structured as interconnected topics which build upon one another to achieve an overall safety vision.

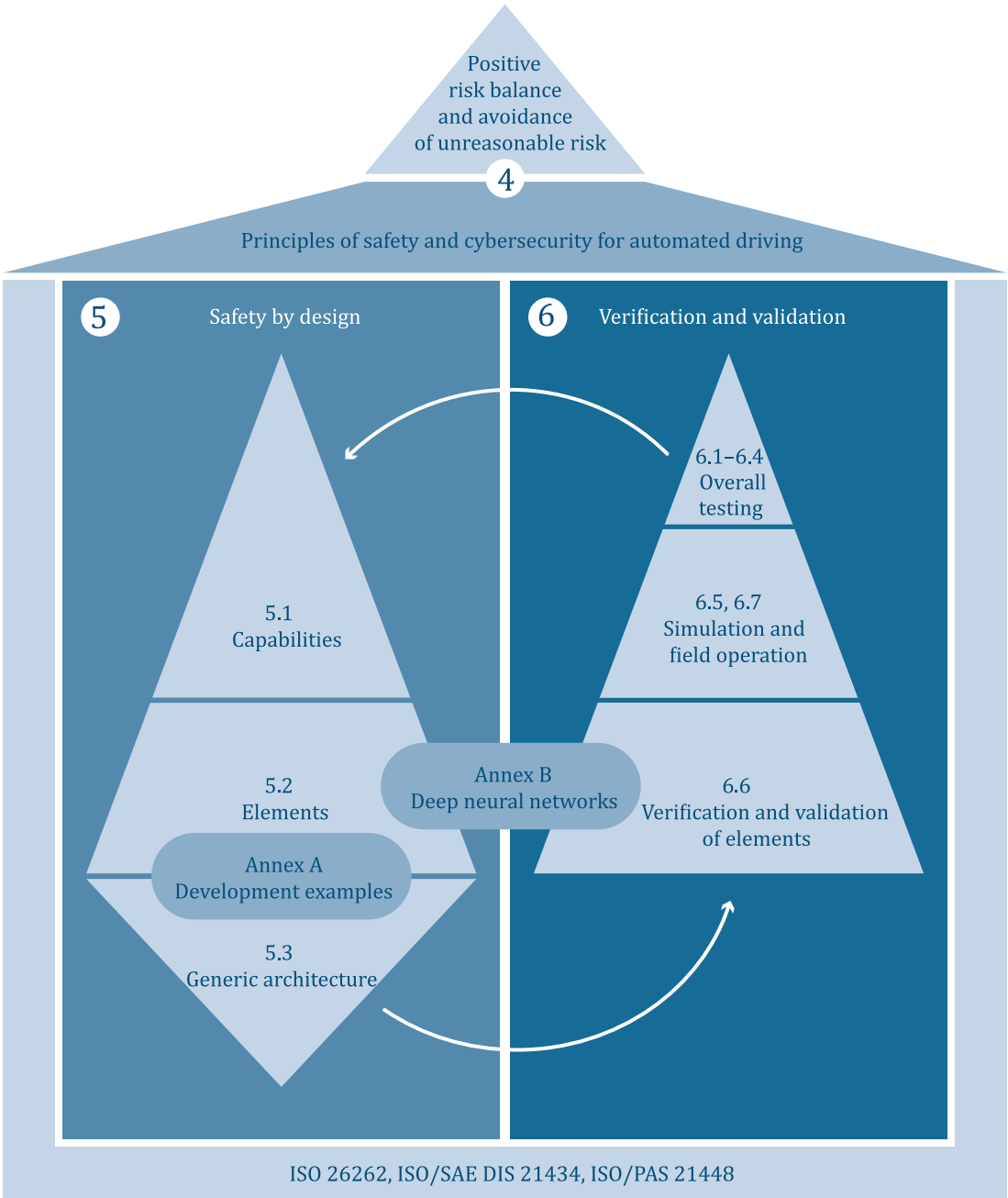

Figure 2 — Structural overview

The structure of this document is summarized in Figure 2. Clause 3 describes the safety and cybersecurity principles recommended to achieve a positive risk balance and to avoid unreasonable risk. Clause 4 describes how measures taken during the system design support the argument for safety and cybersecurity and Clause 5 describes the contribution of evidence provided by the verification and validation activities. The first pillar introduces the three domains for automated driving: safety of the intended functionality (5.2.2, based on ISO/PAS 21448), functional safety (5.2.3, based on the

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

ISO 26262:2018 series) and automotive cybersecurity (5.2.4, based on ISO/SAE 21434[1]). Capabilities for automated driving are then derived from the principles of safety and cybersecurity for automated driving systems and the three dependability domains. The architecture of the development examples (described in Annex A) forms the last element of this pillar.

The second pillar begins by introducing the approach in 6.2 to 6.4 before discussing the quantity of testing (6.5) and simulation (6.6). 6.7 presents the verification and validation of the elements introduced in 5.3. The final block of the second pillar comprises the discussion relating to field operation in 6.8. Both pillars are linked together via the verification and validation approach outlined throughout Clause 6, which combines safety by design and testing with the main strategies applied in verification and validation to solve the challenges discussed throughout this document. Finally, Annex B discusses the use of deep neural networks (DNNs) to realize safety-relevant elements for automated driving.

Various methods are used to aid the reader of this document. This document also uses the following four development examples and their pictograms throughout for further clarity.

— L3 traffic jam chauffeur system (TJCS) as an option for vehicle customers: vigilant conventional driver with driver's license, driving only on structurally separated roads, typically less pedestrians or cyclists compared to other roads, 60 km/h max., only with leading vehicles, no lane changing, no construction sites, only during daylight, without rain, only temperatures higher than freezing point.

— L3 motorway chauffeur system (MCS) as an option for vehicle customers: vigilant conventional driver with driver's license, driving only on structurally separated roads, 130 km/h max., with and without leading vehicles, lane changing, construction sites, at night and during daylight, moderate rain and snow.

— L4 urban chauffeur system (UCS) in fleet operation in urban areas: non-vigilant conventional driver, not capable of driving, no driver's license necessary, 70 km/h max., large ODD with safety driver, very limited ODD without safety driver, allows for teleoperation if necessary.

— L4 automated valet parking system (AVPS) as an option for vehicle customers and in fleet operation: driverless movement within certified parking structures or areas (no vigilant conventional driver, no driver's license necessary), 10 km/h max., ODD focus on off-street parking and logistic areas, scalable use of infrastructure (infrastructure not mandatory but possible up to teleoperation).

4.4 Safety vision

4.4.1 Background

According to the German traffic accident statistics published by the Federal Statistical Office of Germany[2] over 98 % of traffic crash are caused, at least in part, by humans, see Reference [2]. Similarly, the US National Highway Traffic Safety Administration (NHTSA) reports that up to 94 % of serious vehicle crashes in the US are caused at least in part by human error, see Reference [3]. Therefore, introducing automated driving poses great potential for reducing the rates of crashes and accidents. However, there are also major challenges in realizing the full safety benefit of automated driving in order to make automated driving systems reasonably safe and achieve the target of the avoidance of unreasonable risk as recommended by ISO/PAS 21448, and a "positive risk balance compared to human driving performance", as one of the recommendations of the German Ethics Commission in June, 2017[4] focusing on level 4 and level 5 systems without excluding level 3 systems.

Taking a deeper look at the statistics published by Reference [2] which also serve as an indicator of human driving performance, it can be argued that human beings are a reasonable factor for traffic safety. There is an average distance of 300 000 km between two crashes of any severity with respect to a lifetime mileage of 700 000 km.

1) Under preparation. Stage at the time of publication: ISO/SAE DIS 21434:2020.

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

4.4.2 Positive risk balance and avoidance of unreasonable risk

Automated driving can improve performance in most situations compared to that of conventional drivers. However, it cannot completely eliminate the risk of accidents or crashes.

Avoidance of unreasonable risk is a major measure to claim an acceptable level of safety. Its evidence is based on the application of a proactive and reactive driving behaviour, avoidance of crashes as much as "practically possible" while reasonably safe and the avoidance of discrimination on basis of any road user-related characteristics. These judgements are typically made on basis of a combination of qualitative and quantitative assessments, and also on an understanding of good engineering practice and existing standards. It can be argued that applying existing standards and engineering state-of-the-art will result in safer products.

A positive risk balance is a second major measure of an ethically acceptable level of safety. The evidence used for a positive risk balance can be derived from traffic accidents statistics to define acceptance criteria. It is important that validation target values are based on conservative traffic accident statistics of the intended target market, covering dense traffic and challenging traffic scenarios. Different validation target values might apply for different markets. In the long term state-of-the-art automated driving systems will be reflected in the traffic accident statistics and will continuously increase the validation target values of new automated driving systems.

The goal of this document is to present a general approach for tackling the risks introduced by automated vehicles. While this general approach is to be interpreted as a baseline for safe automated driving, it does not define a specific product that is complete and safe.

4.4.3 Principles of safety and cybersecurity for automated driving

The general approach of this document is based on the principles of safety and cybersecurity (PSC) for automated driving systems. They are derived from the considerations of the positive risk balance and the avoidance of unreasonable risk as presented above. Additionally, they are comprising a collection of publications and recommendations from the main public authorities or consumer associations (see Annex C and References [3]-[8]). These principles provide a foundation for deriving a baseline for the overall safety requirements and activities necessary for the different automated driving functions.

The principles of safety and cybersecurity for automated driving are clustered in three groups, each with a dedicated common area of influence (see Figure 3).

The first group named "automated vehicle and related aspects" is addressing overarching aspects on vehicle level including the automated driving system and human factors. Nevertheless, these aspects are essential to achieve overall safety and cybersecurity. This covers technical features as well as process related aspects.

The principles of the second group named "automated driving system" focus on the main technical aspects of the system. These principles constitute functionality related to safety and cybersecurity of the automated driving system itself.

The third group is named "human factors" and is addressing all safety and secure interaction aspects of the ADS and the user. This includes the role of user, clear repartition of role between the user and the ADS as well as takeover scenarios in both directions.

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

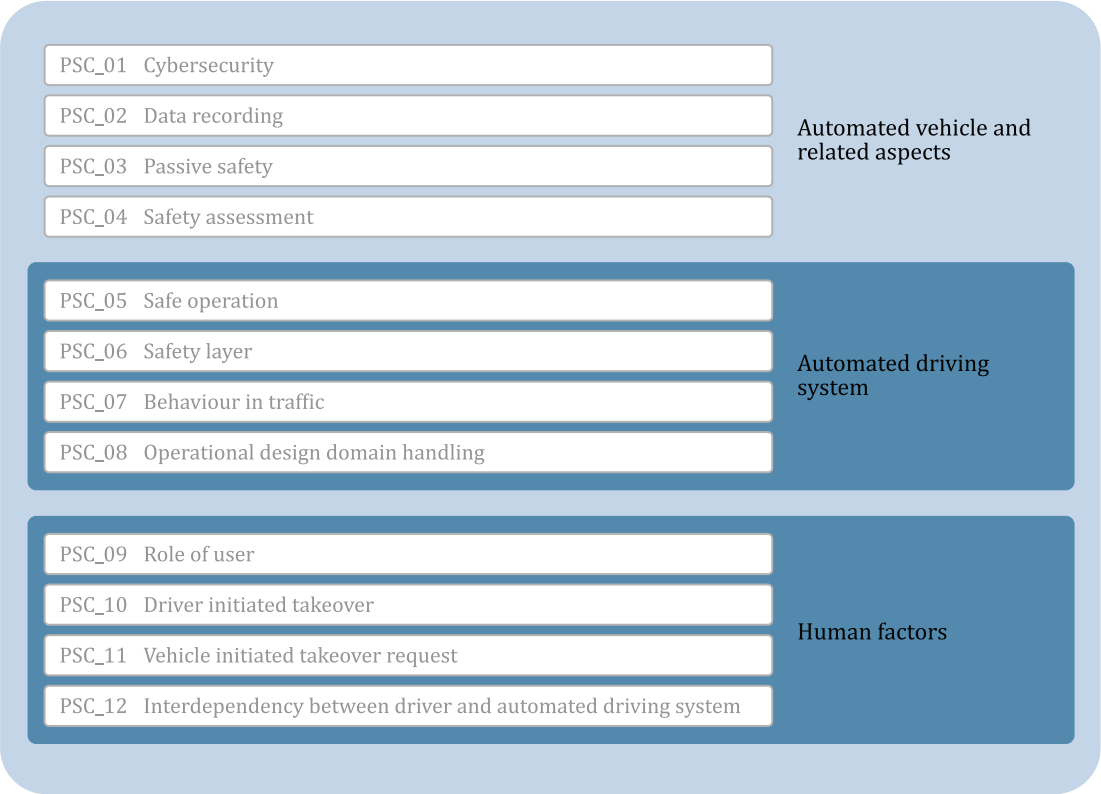

Figure 3 — Groups of principles of safety and cybersecurity for automated driving

The purpose of this document is to highlight safety and cybersecurity-relevant aspects of developing, producing, operating and maintaining automated driving systems the combination of which leads to a safe product on the road. The aspects brought forward contribute toward a foundation for the safety and cybersecurity of automated vehicles.

4.4.3.1 PSC_01: Cybersecurity

When providing an automated driving system, steps are taken to protect the automated driving system from cybersecurity threats.

4.4.3.2 PSC_02: Data recording

Automated vehicles record the relevant data about the status of the automated driving system when a crash is recognized in a manner that complies with the applicable data privacy laws. Data recording must be sufficient to support forensic determination of crash root cause.

4.4.3.3 PSC_03: Passive safety

Crash scenarios

If existing, the vehicle layout considers collision scenarios resulting from vehicle automation.

Alternative seating positions

Occupant protection is considered even when the customer has new uses for the interior that are made possible through the automated vehicle.

NOTE Passive safety is not further detailed in this document as it is fully covered by other standards and regulations.

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

4.4.3.4 PSC_04: Safety assessment

Verification and validation ensure that the safety requirements (ISO 26262 series and ISO/PAS 21448) and safety relevant security requirements (ISO/SAE 21434) are met.

4.4.3.5 PSC_05: Safe operation

Dealing with degradation

If safety-related functions or system components become hazardous (e.g. unavailable, unreliable), the automated driving system:

— is capable of compensating and transferring the automated driving system to a safe condition/state (without unreasonable risk); and

— ensures a safe transition of control to the driver.

Fail-operational (limited to the safety-related function or component)

The loss of safety-related functions or system components do not lead to an unsafe situation.

4.4.3.6 PSC_06: Safety layer

The automated driving system recognizes system limits, especially those that do not allow a safe transition of control to the driver, preventing him or her from reacting to and minimizing the risk. The recognition is performed at least before any driving-cycle or activation of the automated driving system functionality.

4.4.3.7 PSC_07: Behaviour in traffic

Manners on the road

It is important that the behaviour of the automated driving function is be easy-to-understand for surrounding other road users, and also predictable and manageable.

Conforming to rules

It is important that the automated driving system complies with all applicable traffic rules. Crash avoidance manoeuvres can be prioritized over traffic rules if these are not endangering other road users. The principle PSC_09 role of user describes the remaining user responsibilities.

4.4.3.8 PSC_08: Operational design domain handling

ODD determination

As soon as the automated driving system recognizes that system limits, that restrict the safe functionality of the automated driving system are approached or reached, the automated driving system is reacting to compensate or to issue a driver takeover request with a time frame sufficient for takeover in reasonably predictable situations and contexts.

Manage typical situations

The automated driving system takes the relevant situations into account that can reasonably be expected in the ODD and address possible risks including other road users.

4.4.3.9 PSC_09: Role of user

To promote safety, it is important that the driver's state (i.e. state of attentiveness, see Reference [9] and 6.4.1) is suitable for a responsible takeover procedure. The automated driving system keeps the user informed about his responsibilities concerning the required tasks. It also informs the respective driver about safety-relevant driving situations in unmanned driving services.

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

Responsibilities

It is important that the aspects of the driving task which remain under the driver's responsibility are clear to the driver before engaging, while operating and disengaging the automated driving system.

Operating mode awareness

It is important that the automated function ensures that the currently active driving mode can be clearly seen or heard at any time. In addition, a change in driving mode is clearly apparent to the user as well.

4.4.3.10 PSC_10: Driver initiated takeover

Engaging and disengaging the automated driving system requires an explicit interaction from the driver, indicating a high confidence of intent.

NOTE The driver takeover is prioritized against the automated driving system.

4.4.3.11 PSC_11: Vehicle initiated takeover request

Minimal risk condition

If the driver does not accept a takeover request, the automated driving system performs a manoeuvre to minimize risk, resulting in a minimal risk condition. This manoeuvre depends on the situation and the current performance of the automated driving system.

Takeover requests

It is important that vehicle-initiated takeovers are clearly understandable and manageable for the driver.

4.4.3.12 PSC_12: Interdependency between driver and automated driving system

The overall evaluation of system safety takes the effects on the driver due to automation into account, even when they occur after the automated driving period has ended but a direct link to the automated driving part of the journey can be drawn.

5 Systematically developing dependability to support safety by design

5.1 General

This clause describes how the three dependability domains, safety of the intended functionality, functional safety and cybersecurity, work together and how to combine them to create a dependable system. The clause begins by introducing each domain and deriving automated driving capabilities from dependability. It then provides elements that can implement these capabilities. Lastly, it combines all elements by introducing a generic logical architecture (see Figure 4).

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

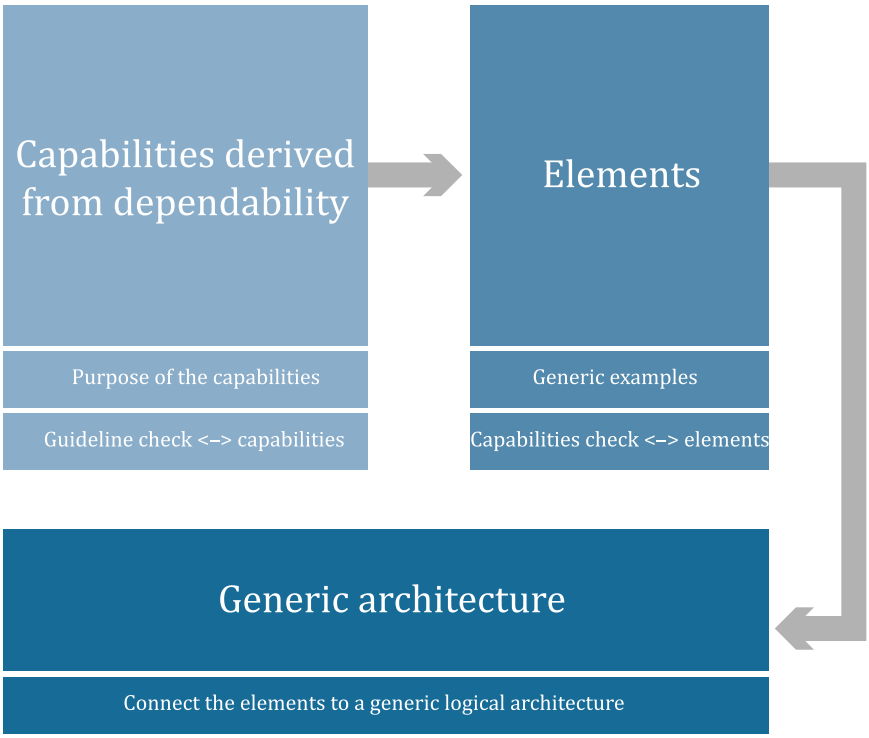

Figure 4 — Systematic development of dependability

5.2 Deriving capabilities of automated driving from dependability domains

Deriving capabilities from dependability domains begins with different worldwide regulatory publications for automated driving systems to identify the requirements that capabilities need to cover in addition to the principles of safety and cybersecurity for automated driving systems. The capabilities cover SOTIF, which includes human factors, functional safety and cybersecurity, which both address the logical and technical architecture. Input requirements are provided from each domain to the others. This subclause provides advice on cybersecurity approaches and measures.

5.2.1 Applying the related safety standards

To ensure a necessary degree of safety, it is important that the system controls safety relevant use cases that result either from the intended use or from unlikely electrical and/or electronic (E/E) faults. The ISO 26262:2018 series and ISO/PAS 21448 use the term foreseeable misuse in the hazard analysis phase of the product development. While performing the foreseeable misuse analysis various actor types are considered which may include both compliant and malevolent vehicle occupants and external actors. The malevolent actor type is one that aligns with both ISO/SAE 21434 and allows for the consideration of intentional misuse or damage to the automated driving system. This is specifically related to systems which relieve the conventional driver from the driving task due to higher automation (SAE L2 and higher). This perspective is partially covered in well-established development standards like the ISO 26262 series and ISO/PAS 21448. This document has the aim for more guidance specifically for the application of the related safety standards for automated driving systems.

Existing standards do not present detailed solutions to some of the most relevant topics of automated driving systems, such as the safety assurance of artificial intelligence (the most relevant algorithms derive from the fields of machine learning and neural networks, see Annex B), human factors and psychology, and the technological capability of the sensory devices used as inputs to the automated driving system. Nevertheless, it is important that safety-related use cases are analysed to ensure the necessary levels of safety. These analyses systematically assess the functional descriptions for possible hazards arising from the intended use and from foreseeable misuse. In addition to a safe design and development process, assessment progresses iteratively from verification to validation and comprises

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

expert appraisals, safety analyses and experiments. Depending on their scope, the different standards support this procedure.

According to ISO/PAS 21448, initially a safe functionality needs to be defined. ISO/PAS 21448 was developed to address the level of risk and hazards caused by the intended functionality, including foreseeable misuse. Danger stemming from E/E malfunctions of the system is addressed by functional safety using the globally established ISO 26262 series, whereby danger as a result of deliberate manipulation is assessed from an ISO/SAE 21434 cybersecurity point of view. Implementing the safety standards ISO/PAS 21448, the ISO 26262 series and ISO/SAE 21434 would allow combining of their procedures and methods. Depending on the development organization's needs, it may be necessary to develop the standards independently, taking their dependencies into account while doing so.

When implementing automated driving system functions, the risks based on the functional and system limits are evaluated. Performance limitations (e.g. based on sensor limitations) could potentially introduce hazardous situations. Development standards will aid developers in managing the complexity of the systems, estimating possible risks and addressing adequate measures.

Finally, availability of safety-related capabilities of a system is an important topic, which is partially covered by the ISO 26262 series.

As stated in ISO/PAS 21448, it becomes increasingly important to approach the problem from an analytical perspective, addressing the safety of the system by designing scenario-based system behaviours and addressing the technological capability including availability aspects through the analysis of use cases and scenarios to design a robust and safe system. This document defines a typical set of arguments and technological capabilities as the concept of safety by design, and this is the approach that is detailed throughout this document.

To achieve the balance between safety with respect to fail-safe behaviour and availability of customer functionality, the design is analysed and built from the top down. The first analysis is carried out irrespective of the generic logical architecture. The process includes risk assessments to determine the safety requirements of the system being designed. Ultimately, this evolves into a safety concept, defining safety mechanisms to support the safety goals.

5.2.2 ISO/PAS 21448 - Safety of the intended functionality

The basic concept of the ISO/PAS 21448 - Safety of the Intended Functionality (SOTIF) approach is to introduce an iterative function development process (see Figure 26) that includes verification and validation. This approach leads to an intended function which can be declared safe by applying progressive improvement during development. Several activities will be derived based on an approach that argues that these activities are adequate for developing a safe automated function. This approach assumes that there is an area of known scenarios with defined behaviour and an unknown area with potential harm. The automotive development goal is to reduce the impact of known potentially harmful scenarios and the amount of unknown scenarios to an acceptable level of residual risk.

Clause 5 can be referred to at this point since the main evidence to demonstrate that the system is reasonably safe enough for road users is provided via verification and validation activities.

5.2.3 ISO 26262 series - Functional safety

In the ISO 26262 series the risk is determined and communicated or mapped using automotive safety integrity levels (ASIL). These levels are representing related requirements. This document will not attempt to directly trace and decompose ASIL values in a traditional application of functional safety, but these may be indirectly inferred from the architectural examples that are derived in the subclauses below.

A starting point in the ISO 26262 series is the risk determination based on a defined function including a first preliminary architecture. Therefore, the approach of ISO/PAS 21448 SOTIF can be used to support – applying functional safety in accordance with the ISO 26262 series – the creation of the item definition. This item definition per the ISO 26262 series needs to include a definition of the functions, including their dependencies on and interaction with the environment and other items/vehicles. Based on the item definition, a hazard analysis and risk assessment can then be carried out. [illegible]

© ISO 2020 – All rights reserved

AVA Documents Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

requirements or safety goals for the function and its related system. The next technical steps are developing functional and technical safety concepts.

The first edition of the ISO 26262:2011 series was created based on the knowledge of state-of-the-art systems in automotive industries (such as steering, braking and airbag systems, etc.) and does not completely address very complex, distributed systems and how to deal with availability requirements. The second edition resolves some of these issues, but further interpretations are needed.

Meeting all challenges will result in the definition of the safe intended functionality. In other words, weaknesses of the specification or the technologies have been considered (SOTIF) and possible E/E faults can be controlled by the system or by other measures (the ISO 26262 series). Consequently, it will be possible to declare the automated system safe without considering cybersecurity issues, which are currently not covered by ISO/PAS 21448 or the ISO 26262:2018 series.

5.2.4 ISO/SAE 21434 - Automotive cybersecurity

Cybersecurity focuses on risks presented by active adversaries in the form of creative, determined and malicious entities acting intentionally. This leads cybersecurity to utilize additional analysis tools and technical mechanisms that nevertheless also affect safety.

The ISO/SAE 21434 describes the tasks for a secure development process, with a special focus on risk analysis. For automated driving systems the ISO/SAE 21434 is therefore fully applicable. However, while using a certain risk approach, it is important to think about specifics of automated driving.

The automotive industry is facing new challenges in automated driving due to the extreme connectivity within automated driving vehicles and between those vehicles and their operating environment. Connectivity additions include new interfaces between the control functions of connected vehicles, IT backend systems, and other external information sources (see Figure 5). This rich attack surface creates considerable interest for malicious actors with various goals. As the connectivity and level of automation increase, the likelihood of an attack will increase. Therefore, the cybersecurity risk will be higher without proper mitigation.

The other dimension of a cybersecurity risk is the impact of an attack. Here, there are also differences between human controlled driver assistance functions and system controlled automated driving functions. As the driver is out of the loop, they are not able to recognize an attack and to react to it. So, the impact is probably higher than in actual released systems. There are approaches to use the safety analysis results as input for cybersecurity analysis, see ISO/SAE 21434:—, Clause 8, RQ-08-07.

Another interesting overlap is between the disciplines of cybersecurity and safety of the intended function. There are some real-world attacks focus on vehicles but without directly attacking them. For example, if the lane marking on the road is changed, or if a misleading traffic sign is projected on a house wall, the vehicle may be deceived into leaving its lane. It is acknowledged that these kinds of attacks are neither solely a cybersecurity nor a SOTIF topic. It is rather a joint effort to account for these topics, probably with additional domains. Thereto, one might use methods from both domains as follows. The risk on objects outside the vehicle is considered by the different standards. It is important that mitigations are evaluated, both inside and outside the vehicle [see ISO/SAE 21434:—, 8.9, RQ-08-12 (risk transfer)]. In the end, it may not be the task for road vehicle cybersecurity to mitigate these risks with typical cybersecurity measures like integrity checks or proofing the authenticity as the attacked environment is not cyber-physically connected to the considered system. Also, for SOTIF, it is hard to cover these scenarios as it deals mainly with incorrect usage of the system itself, instead of the environment. Nevertheless, these attacks and risks are considered when defining the ODD. Resulting measures to cover these risks have then also to be covered by means of validation.

In short, we have advanced to a level where vehicles cannot maintain a safe state unless they also operate securely. Most importantly, cybersecurity processes of ISO/SAE 21434 are applied as a starting point to ensure that for example attackers cannot gain arbitrary control of a vehicle's movement, and that attacks are exceptionally difficult to scale to the point of simultaneously exploiting multiple vehicles.

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

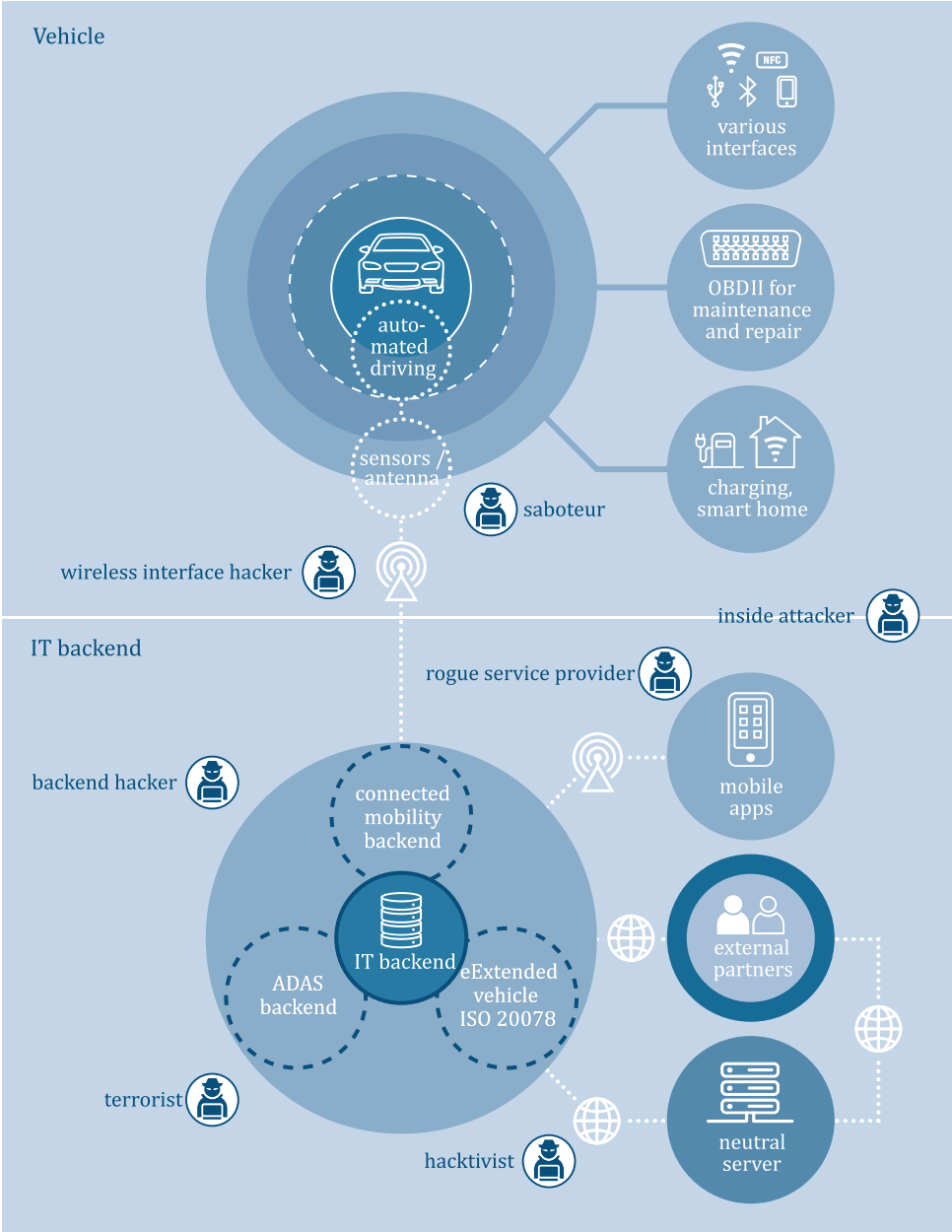

Figure 5 — Automotive cybersecurity

5.2.5 Capabilities of automated driving

5.2.5.1 Initial derivation of capabilities

An automated driving system that complies with the principles outlined in 4.4.3 has a basic set of system properties that are specified here as capabilities. The following will discuss the capabilities present when the overall system is considered safe.

The capabilities are divided into fail-safe capabilities (FS) and fail-degraded capabilities (FD). Fail-safe capabilities provide and enable customer value, but they can also bring the system to achieve a minimal risk condition in the event of failure. Fail-safe capabilities can be discontinued, because the safety relevance of their unavailability is low enough or is covered by the fail-degraded capabilities.

During the time in which the system is performing nominally, system operation may be understood using the classic sense – plan – act design paradigm from robotics and automation literature. In this

© ISO 2020 – All rights reserved

AVA Documents Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

model, sensing and perception (including localization), planning and control, and actuation and stability provide a general, implementation-independent view of the automated driving system. Figure 6 illustrates this general model:

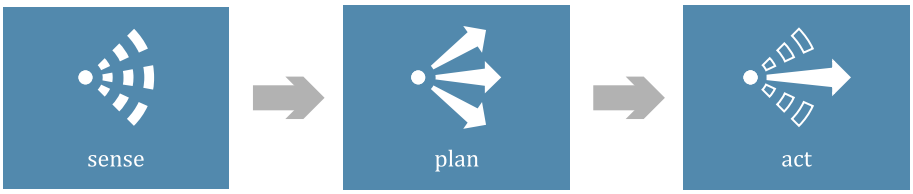

Figure 6 — Sense – plan – act design paradigm

Based on the allocation of capabilities to the basic functions for sense – plan – act, it is possible to allocate requirements for elements that the automated vehicle is reasonably safe as depicted in Figure 7.

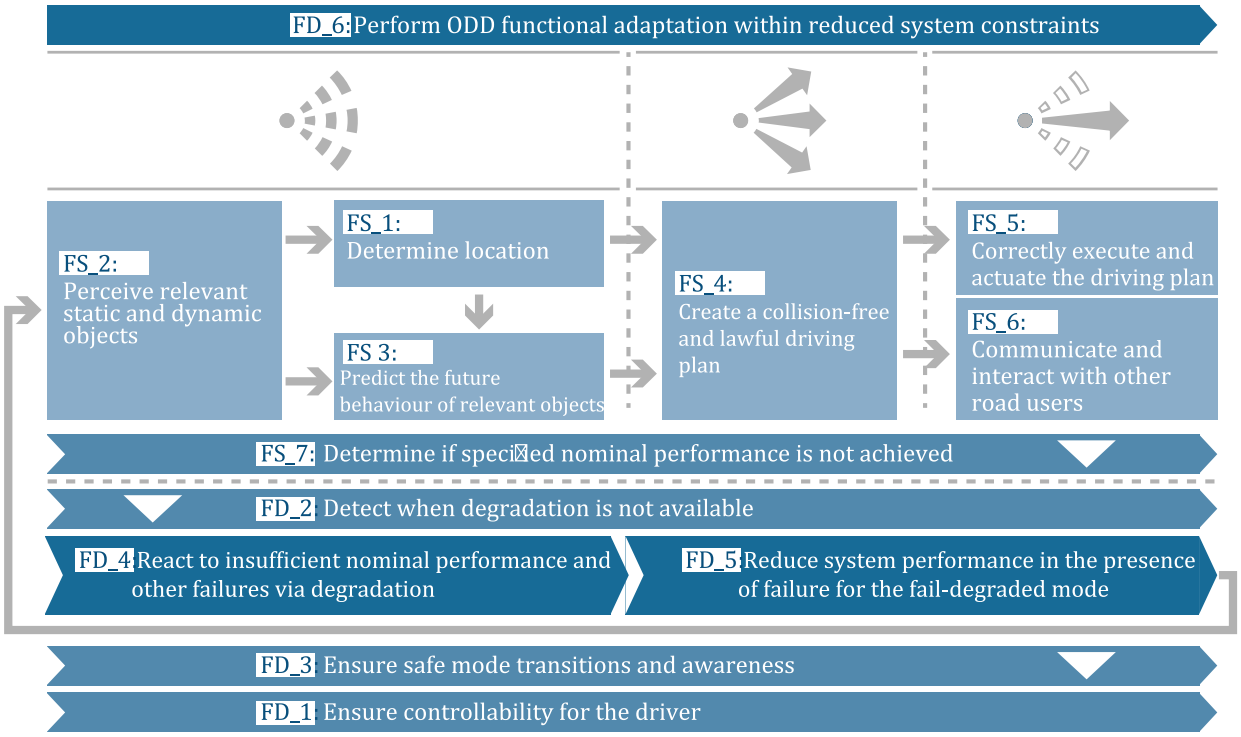

Figure 7 — Realizing fail-safe and fail-degraded capabilities

5.2.5.2 Overview of the capabilities

FS_1: Determine location

The system is able to determine its location in relation to the ODD. The vehicle is able to decide if it is inside or outside of a location-specific ODD. The location in the ODD may be required, depending on the item definition.

NOTE The location is stored as initial situation of any realization of capabilities, including the relevant perceived proximity information.

FS_2: Perceive relevant static and dynamic objects

All entities that an automated driving system requires for its functional behaviour are perceived, optionally pre-processed, and provided correctly. The highest priorities are placed on entities with an associated risk of collision, also confirming their plausibility. Sample entities include dynamic objects

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

(e.g. other road users and characteristics of the respective movement), static instances (e.g. road boundaries, traffic guidance and communication signals) and obstacles.

FS_3: Predict the future behaviour of relevant objects

The relevant environment model is extended by the predicted future state. The aim is to create a forecast of the environment. The intention of the relevant objects is interpreted in order to form the basis for predicting future motion.

FS_4: Create a collision-free and lawful driving plan

To ensure a collision-free and lawful driving policy, the following is respected.

Maintaining a safe lateral and longitudinal distance to other objects:

— complying with all applicable traffic rules whenever and wherever the automated vehicle is driving;

— considering potential areas where objects, other road users, or animals may be occluded;

— in unclear situations the right of way is given, not taken;

— if a crash can be avoided without endangering other road users, traffic rules may be compromised if necessary, to avoid injury or preserve life;

— considering the inadequacy and uncertainty of other road users and automated vehicles.

FS_5: Correctly execute and actuate the driving plan

The corresponding actuation signals for lateral and longitudinal control are generated based on the driving plan.

FS_6: Communicate and interact with other road users

For safety-related use cases and depending on the ODD, automated vehicles communicate and interact with other road users.

FS_7: Determine if specified nominal performance is not achieved

Any element of the automated driving system can, either on its own or in combination with others, result in adverse behaviour. Therefore, the adverse nominal performance of the system or violation of ODD conditions are detected by appropriate mechanisms. FD_4 covers the reaction to detected adverse behaviour.

Typical aspects for influencing the nominal performance are:

— unwanted human factors, including foreseeable misuse and manipulations;

— deviation of the intended functionality;

— technological limitations;

— environmental conditions;

— systematic and random failures.

Capabilities for recovering to nominal performance, as specifically defined herein, are possible but are not considered further in this document, because they have no direct safety relevance. Fulfilling the capabilities is necessary but not sufficient for safe system operation. Additional capabilities will be required depending on the specified functionality and product.

FD_1: Ensure controllability for the driver

The vehicle operator's level of control varies depending on the automation level as per SAE J3016:2018 and the use case definition and therefore is ensured.

Documents
Document Sharing Hub

© ISO 2020 – All rights reserved

Full document details: PD ISO/TR 4804:2020

FD_2: Detect when degradation is not available

It is ensured that a possible unavailability of the fail-degraded capability is detected. If the degradation strategies depend on the degradation reason, the degradation reason is identified.

FD_3: Ensure safe mode transitions and operating mode awareness

Transitions of driving mode are performed correctly and controlled by the affected vehicle operator if necessary. It is important that the affected driver is aware of the current driving mode and their responsibility deriving from it.

For example, actuating an automated mode is permitted only when inside the ODD, and it will be deactivated with the vehicle operator taking control again or misusing the system and it will be deactivated prior to leaving the ODD.

FD_4: React to insufficient nominal performance and other failures via degradation

Due to possibly unavailable nominal performance capabilities and other failures (e.g. based on hardware faults), it is important that the system degrades within a well-defined amount of time.

FD_5: Reduce system performance in the presence of failure for the fail-degraded mode

The reaction in case of failures during fail-degraded mode is defined.

FD_6: Perform ODD functional adaptation within reduced system constraints

Automated driving system operation with ODD functional adaptation is actuated as nominal capabilities with new limits. Multiple functional adaptations are possible. The new limits are defined such that the functional adaptations are safe. If suitable, a safe functional adaptation can be a non-permanent operation with a defined time frame for an additional required reaction.

The selection matrix in Table 2 is an example to demonstrates the state of completeness of the derived capabilities to provide evidence of traceability to the principles from 4.4.3. Such kind of table needs to be adapted for a concrete development project.

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

Table 2 — Example of a selection matrix for the traceability of the capabilities

ID	**PSC_01** Cybersecurity	**PSC_02** Data recording	**PCS_03** Passive safety	**PSC_04** Safety assessment	**PSC_05** Safe operation	**PSC_06** Safety layer	**PSC_07** Behaviour in traffic	**PSC_08** Operational design domain handling	**PSC_09** Role of user	**PSC_10** Driver initiated takeover	**PSC_11** Vehicle initiated takeover request	**PSC_12** Interdependency between driver and ADS
FS_1 Determine location	×			×			×	×				
FS_2 Perceive relevant static and dynamic objects	×			×			×					
FS_3 Predict the future behaviour of relevant objects	×			×			×					
FS_4 Create a collision-free and lawful driving plan	×			×			×					
FS_5 Correctly execute and actuate the driving plan	×			×			×					
FS_6 Communicate and interact with other road users	×			×			×					
FS_7 Determine if specified nominal performance is not achieved	×			×		×		×				
FD_1 Ensure controllability for the driver	×			×	×				×	×	×	×
FD_2 Detect when degradation is not available	×			×	×							
FD_3 Ensure safe mode transitions and operating mode awareness	×			×	×	×			×	×	×	×
FD_4 React to insufficient nominal performance and other failures via degradation	×			×	×	×						
FD_5 Reduce system performance in the presence of failure for the fail-degraded mode	×			×	×	×						
FD_6 Perform ODD functional adaptation within reduced system constraints	×			×	×	×		×			×	

As shown in Table 2, the safety principle for safety assessment traces to all capabilities. This derives from the expectation that product development would be responsible for delivering a necessary level of evidence for the verification and validation of the capabilities, which may then be reviewed by an assessing group. While this is recognized here, it will be discussed in greater detail in later subclauses where the full logic and rationale surrounding the methodology for validating an automated driving system is developed.

5.2.6 Minimal risk conditions and minimal risk manoeuvres

A minimal risk manoeuvre (MRM) is the system's capability of transitioning the vehicle between minimal risk conditions (MRC). The concept of MRCs and MRMs are derived considering safety at vehicle level and are defined in accordance with the principles of the ISO 26262 series as an operating mode (in the case of a failure) of an item with a tolerable level of risk. In terms of ISO 26262-1:2018, an MRM is an emergency operation to reach an MRC – referred to as a safe state. This document expands the definition to also include fail-degraded mode and takeovers by the vehicle operator. Final MRCs refer to MRCs that allow complete deactivation of the automated driving system, e.g. standstill or takeover by the vehicle operator. Figure 8 visualizes this general principle.

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

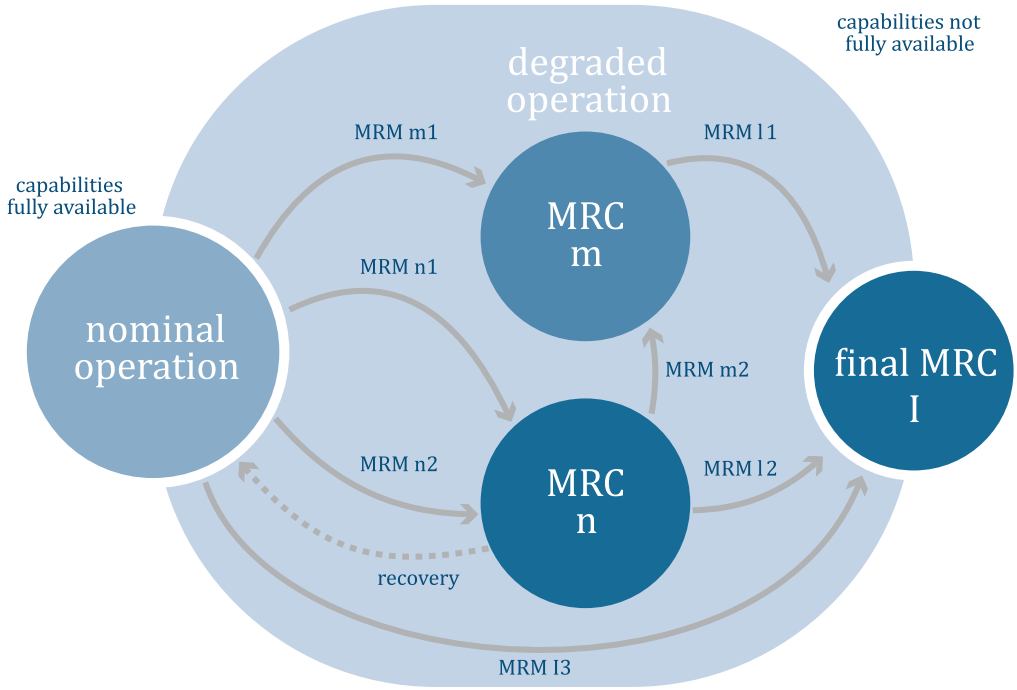

Figure 8 — Minimal risk manoeuvre and minimal risk condition

The purpose of the MRM is to transfer the vehicle to a minimal risk condition. Due to the complexity of automated systems and risk-influencing conditions, several MRCs and MRMs could be conducted consecutively. If not all fail-safe capabilities are available, the system will be in fail-degraded mode and the remaining fail-degraded capabilities will reach a minimal risk condition by executing an appropriate minimal risk manoeuvre. The fail-degraded mode is a time-limited operational domain in which the frequency of its occurrence needs to be reduced whenever possible. The acceptable time for the fail-degraded mode depends on the remaining capabilities in the current system. One principle concept of the ISO 26262 series is that every order of magnitude in reducing the frequency of possible harm lowers the required safety integrity level of the automated driving system. Table 3 defines the conditions between which the MRM will allow for safe transition. Here, "Nominal Operation" is handled like an MRC. Specific examples are outlined in Annex A.

Table 3 — Minimal risk conditions

ID	MRC	Definition	Possible reasons
MRC_1	Takeover by the driver	The driver has completely taken over the driving task.	— Known future limitation in ODD — Limitations or the driver (if present) has initiated takeover — It is detected that degraded performance is not available (FD_2)
MRC_2	Limited operation	Vehicle is still operational within reduced capabilities. There could be several limited operation conditions depending on the functional definition and remaining capabilities.	— Deviations from nominal state, reduced capabilities
MRC_3	End of operation	This condition describes a vehicle state that allows safe deactivation of the function.	— Severe system failures, loss of capabilities, missing driver takeover

© ISO 2020 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

The following MRMs are proposed:

Table 4 — Proposed MRMs

ID	MRM	Definition	Target condition
DTO	Driver takeover request	Request takeover by the driver.	MRC_1 Takeover by the driver
MRM_2	Limit function state	Transition to limited operation. Depending on the MRC and the actual state, several MRM variants are possible.	MRC_2 Limited operation
MRM_3	Comfort stop	Comfortable transition to end of operation.	MRC_3 End of operation
MRM_4	Safe stop	Due to severe failure, next possible stop when suitable according to driving situation and surroundings to end operation.	MRC_3 End of operation
MRM_5	Emergency stop	Due to severe failure, immediate stop is necessary to avoid unreasonable risk.	MRC_3 End of operation
Recovery	Recovery	Limitations of capabilities are resolved and therefore nominal state is reached again.	Nominal state

Using the list of MRMs in Table 4, the potential failure modes are reflected within the overall system. It is important to apply several analysis methods from related safety standards (e.g. failure analysis techniques such as FMEA or FTA, analysis for the intended use of the system and its misuse). The outcome of such analysis measures is to define all desired safe states for each component and to characterize how these safe states enable the MRCs and MRMs of the integrated system.

5.3 Elements for implementing the capabilities

In addition to the different SAE levels and ODD definitions, there are many possibilities when implementing automated driving. The development examples introduced in Annex A can be considered possible real-world implementations that fulfil the capabilities introduced in 5.2.2. This subclause provides guidance about fulfilling the fail-safe and fail-degraded capabilities with real-world elements generically, but in a manner to ensure a reasonable level of safety. 5.3.1 discusses possible implementations of each capability by introducing elements. 5.3.2 then discusses each element in detail.

5.3.1 Implementing the capabilities

The following describes the capabilities in relation to their associated elements. Elements in light blue are for the sake of clarity displayed without further inputs. Each of those elements is represented as a whole with all relevant inputs in one figure only.

5.3.1.1 FS_1: Determine location

Location information can be used to select correct lane (e.g. for making turns) or to know which local traffic rules apply. In addition to that it is ensured that the automated driving system operates only within specified system limits if the intended use of the automated driving system is restricted. Therefore, it is important that the automated driving vehicle is located adequately, using fused environmental sensor information from sensor fusion algorithms (see 5.3.1.4). To achieve appropriate localization, it may be necessary to link additional a-priori information from outside of the onboard perception's performance (e.g. via map information, referencing detected events to a unique coordinate system), either in range or in interpretation, with the automated driving vehicle. Localization may consider information from egomotion to predict whether the automated vehicle is about to exceed an ODD limit. This is a precondition for activating the automated driving system, which prevents operators from misusing the automated driving system outside the ODD (see Figure 9).

© ISO 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 4804:2020

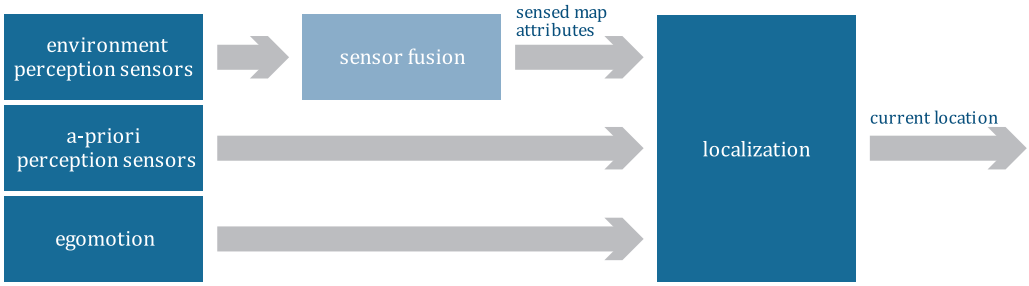

Figure 9 — FS_1: Determine location

5.3.1.2 FS_2: Perceive relevant static and dynamic objects

All entities that an automated driving system requires to account for its functional behaviour are perceived, pre-processed and provided safely. The highest priority is placed on entities with the highest associated risk of collision. Example entities include dynamic instances (e.g. other road users and characteristics of the respective movement), static instances (e.g. road boundaries, traffic guidance signals), and obstacles exceeding a critical size. The main element for this is the sensor fusion element, where the different inputs from onboard sensors, localization, egomotion and optional V2X information are used to generate the present world model. Traffic rules might be used to optimize the content of the world model. The semantic knowledge of the perceived environment is important for later interpretation, prediction and planning. Consequently, the automated driving system knows what has been detected and where. The definition of relevant objects depends on the item definition.

The detection of static instances can be supported by the localization element, which provides a map and a location of the automated driving vehicle on the map. With entities marked on the map, this supports the sensor fusion algorithm with another independent sensor (see Figure 10).

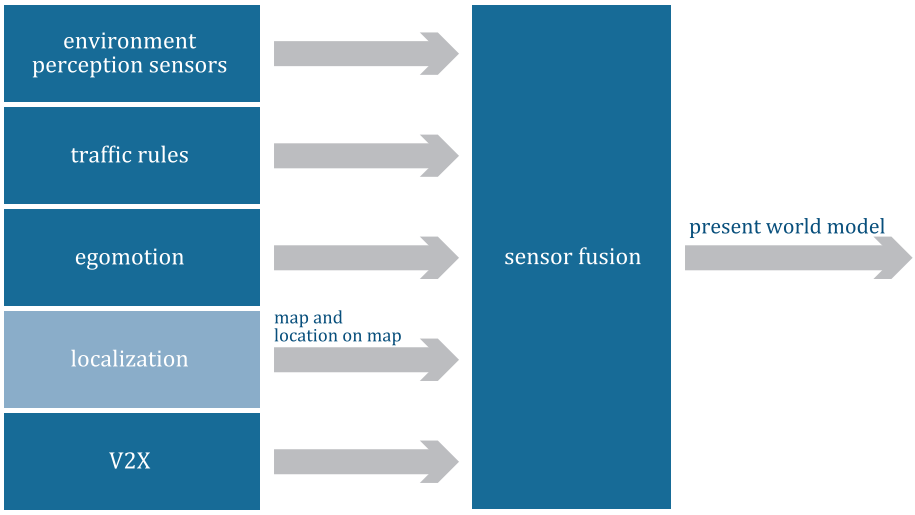

Figure 10 — FS_2: Perceive relevant static and dynamic objects

5.3.1.3 FS_3: Predict the future behaviour of relevant objects

The present world model recorded as the output of FS_2 (5.3.1.2) may not suffice as an input for the safe and lawful creation of a driving plan FS_4 (5.3.1.4). Therefore, it is important to extended it to reflect not only the current but also the predicted future state of the world model in order to generate a complete description of the dynamic driving situation or "scene" (see Figure 11). It also considers the intention of other dynamic objects and obstacles including those that are partially occluded as the basis for predicting future motions. Additionally, current environment conditions such as low road friction and reduced sensor performance (e.g. fog, mist, heavy rain) are taken into consideration for

© ISO 2020 – All rights reserved
Documents Document Sharing Hub
Full document details: PD ISO/TR 4804:2020