ISO/DPAS 8800

ISO/TC 22/SC 32

~~ISO/CD PAS 8800(en)~~

Secretariat: ~~–~~JISC

Date: 2024-08-06

Road vehicles — Safety and artificial intelligence

Véhicules routiers — Sécurité et intelligence artificielle

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

Published in Switzerland

Contents

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 document 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).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at ~~www.iso.org/patents.~~www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

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~~www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 22, *Road vehicles*, Subcommittee SC 32, *Electrical and electronic components and general system aspects*.

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.

Introduction

The purpose of this document is to provide industry-specific guidance on the use of AI systems in safety-related functions. It is not restricted to specific AI methods or specific vehicle functions.

This document defines a framework for managing AI safety that tailors or extends existing approaches currently defined in the ISO 26262 series and in ISO 21448.

Functional safety-related risks associated with malfunctioning behaviour of an AI system are addressed by tailoring or extending relevant clauses from ISO 26262-4, ISO 26262-6 and ISO 26262-8.

The risks related to functional insufficiencies in the AI system are addressed by extending the concepts and guidance provided by ISO 21448. A causal model for understanding the sources of functional insufficiencies in the AI system is proposed. The model is used to derive a set of safety requirements on the AI system as well as a set of risk reduction measures.

NOTE 1 ISO 21448 is applicable to intended functionalities where proper situational awareness is essential to safety and where such situational awareness is derived from sensors and processing algorithms, especially functionalities of emergency intervention systems and systems with ISO/SAE PAS 22736 levels 1 to 5 for driving automation. It is therefore possible that systems utilize AI technologies that do not fall within the scope of ISO 21448.

EXAMPLE 1 ISO 21448 does not apply to the development of an engine control unit that uses AI to optimize its performance whereas this document does.

This document recognizes that due to the wide range of applications of AI and associated safety requirements, as well as the rapidly evolving state-of-the-art, it is not possible to provide detailed requirements on the process or product characteristics required to achieve an acceptably low level of residual risk associated with the use of AI systems. Therefore, in addition to providing guidance for tailoring or extending the ISO 26262 series and ISO 21448, this document focuses on the principles that support the creation of a project-specific assurance argument for the safety of the AI elements within on-board vehicle systems. This includes proposing risk reduction measures during the design and operation phases using an iterative approach to reducing risk as outlined in ISO/IEC Guide 51.

Hazard analysis and risk analysis are beyond the scope of this document. These are considered a part of the vehicle level systems safety engineering activities described in the ISO 26262 series and ISO 21448, or in application of specific standards such as ISO TS 5083.

ISO/IEC TR 5469 provides generic guidance for the application of AI technologies as part of safety functions, independent of specific industry sectors. Many of the concepts outlined in ISO/IEC TR 5469 can be applied in the context of road vehicles. There is therefore a close relationship to concepts described within this document and ISO/IEC TR 5469.

ISO/IEC TR 5469 provides classification schemes to determine the safety requirements on the AI/ML function. These include the usage level and AI technology class.

The usage level is related to the nature of the task being performed by the engineered AI system.

NOTE 2 The usage levels are described in ISO/IEC TR 5469:2024, 6.2.

The technology class is related to the problem complexity and the transferability of existing standards to demonstrating an adequate level of safety based on properties of the target function and the AI technology used.

NOTE 3 For the technology classes, see ISO/IEC TR 5469:2024, 6.2.

This document does not explicitly call out the classes and usage levels of ISO/IEC TR 5469.

EXAMPLE 2 For some AI technology, the application of ISO 26262 is deemed to be sufficient. This corresponds to Class I of ISO/IEC TR 5469.

The guidance outlined within this document is relevant for all usage of AI for which safety requirements can foreseeably be allocated either through:

a) the use of AI for the functionality itself;

b) the use of AI as a safety mechanism.

NOTE 4 These usages correspond to the usage levels A1, A2, C of ISO/IEC TR 5469. In all cases, the applicability of the guidance provided within this document can be determined by the allocation of safety requirements to the AI technology, whereas the usage levels of ISO/IEC TR 5469 can be used to support the requirements elicitation process.

This document is aligned with standards and documents developed by ISO/IEC JTC1/SC42. AI-specific definitions are used from ISO/IEC 22989, unless in conflict with safety-specific definitions.

Other documents developed within ISO/IEC JTC1/SC42 can be used to provide additional guidance on specific aspects of AI that are relevant to safety-related properties. Examples of such documents include ISO/IEC TR 24027 and ISO/IEC TR 24029-1.

This document harmonizes the concepts already described in ISO 21448:2022, Annex D.2 and ISO/TS 5083:~~2024,~~20—[1], Annex B whilst extending these with specific guidance regarding the definition of safety requirements of machine learning (ML), ML safety analyses and the creation of associated safety evidence during the development and deployment lifecycle.

ISO/TS 5083~~:~~20—, Annex B is an application of this document to automated driving systems (ADS).

The relationship with the above-mentioned documents is summarized in Table 1–1.

Field Code Changed

Table 1–1 — How this document relates to other publications on AI safety

Publication	Relationship with this document
ISO/IEC 22989	AI specific definitions are used from ISO/IEC 22989, unless in conflict with safety-specific definitions. Safety-related properties are a subset of generic AI properties described in ISO/IEC 22989.
ISO/IEC TR 5469	This document does not explicitly call out the classes and usage levels of ISO/IEC TR 5469. This document considers and adapts to road vehicles the general framework described in ISO/IEC TR 5469 on safety properties, virtual testing and physical testing, confidence in use of AI development frameworks and architectural redundancy patterns.
ISO 26262	This document is a tailoring or extension of ISO 26262 for AI elements of the system. See ~~Clause 5~~Clause 5 for details.
ISO 21448	This document is a tailoring or extension of ISO 21448 for AI elements of the system. See ~~Clause 5~~Clause 5 for details.

[1] Under preparation. Stage at the time of publication: ISO/DTS 5083.

Publication	Relationship with this document
ISO TS 5083:~~2024~~20—	ISO TS 5083:~~20—~~, Annex B2 is an application of this document to automated driving systems (ADS).

This document adds the following contents with respect to the documents listed in ~~Table 1-1:~~Table 1-1:

— tailoring or extensions of ISO 26262 and ISO 21448 required specifically for AI elements of the system (referred to as AI systems);

— a conceptual model for reasoning about errors and their causes specific to AI systems;

— a reference AI safety lifecycle;

— the safety assurance argument for AI systems;

— a method for deriving AI safety requirements for AI systems;

— considerations for the design of safe AI systems;

— considerations on data management for the AI systems;

— a verification and validation strategy for AI systems;

— a safety analysis approach for AI systems (focused on insufficiencies);

— activities during operation required to ensure the continuous AI safety.