

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*

iTeh Standards  
(<https://standards.iteh.ai>)  
Document Preview

ISO/DPAS 8800

<https://standards.iteh.ai/catalog/standards/iso/f16a69d8-8442-461c-a36b-5756f4e356a7/iso-dpas-8800>

ISO/~~CD-PASDPAS~~ 8800:2024:(en)

© ISO 2024

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  
E-mail: [copyright@iso.org](mailto:copyright@iso.org)  
Website: [www.iso.org](http://www.iso.org)

Published in Switzerland

iTeh Standards  
(<https://standards.iteh.ai>)  
Document Preview

ISO/DPAS 8800

<https://standards.iteh.ai/catalog/standards/iso/fl6a69d8-8442-461c-a36b-5756f4e356a7/iso-dpas-8800>

Contents

Foreword ..... viii

Introduction ..... ix

1 — Scope ..... 1

2 — Normative references ..... 1

3 — Terms and definitions ..... 2

3.1 — General AI-related definitions ..... 2

3.2 — Data-related definitions ..... 8

3.3 — General safety-related definitions ..... 9

3.4 — Safety: Root cause, error and failure-related definitions ..... 12

3.5 — Miscellaneous definitions ..... 13

4 — Abbreviated terms ..... 15

5 — Requirements for conformity ..... 16

5.1 — Purpose ..... 16

5.2 — General requirements ..... 16

6 — AI within the context of road vehicles system safety engineering and basic concepts ..... 17

6.1 — Application of the ISO 26262-series for the development of AI systems ..... 17

6.2 — Interactions with encompassing system-level safety activities ..... 18

6.3 — Mapping of abstraction layers between ISO 26262, ISO/IEC 22989 and this document ..... 22

6.4 — Example architecture for an AI system ..... 25

6.5 — Types of AI models ..... 26

6.6 — AI technologies of a ML model ..... 26

6.7 — Error concepts, fault models and causal models ..... 27

6.7.1 — Cause-and-effect chain ..... 27

6.7.2 — Root cause classes ..... 28

6.7.3 — Error classification based on the safety impact ..... 30

7 — AI safety management ..... 31

7.1 — Objectives ..... 31

7.2 — Prerequisites and supporting information ..... 31

7.3 — General requirements ..... 31

7.4 — Reference AI safety life cycle ..... 34

7.5 — Iterative development paradigms for AI systems ..... 1

7.6 — Work products ..... 3

8 — Assurance arguments for AI systems ..... 3

8.1 — Objectives ..... 3

8.2 — Prerequisites and supporting information ..... 3

8.3 — General requirements ..... 4

8.4 — AI system-specific considerations in assurance arguments ..... 4

8.5 — Structuring assurance arguments for AI systems ..... 6

8.5.1 — Context of the assurance argument ..... 6

8.5.2 — Categories of evidence ..... 7

8.6 — The role of quantitative targets and qualitative arguments ..... 8

8.7 — Evaluation of the assurance argument ..... 9

8.8 — Work products ..... 10

9 — Derivation of AI safety requirements ..... 10

9.1 — Objectives ..... 10

9.2 — Prerequisites and supporting information ..... 10

9.3 — General requirements ..... 11

9.4	General workflow for deriving safety requirements	12
9.5	Deriving AI safety requirements on supervised machine learning	15
9.5.1	The need for refined AI safety requirements	15
9.5.2	Derivation of refined AI safety requirements to manage uncertainty	16
9.5.3	Refinement of the input space definition for AI safety lifecycle	19
9.5.4	Restricting the occurrence of AI output insufficiencies	19
9.5.5	Metrics, measurements and threshold design	23
9.5.6	Considerations for deriving safety requirements	24
9.6	Work products	24
10	Selection of AI technologies, architectural and development measures	25
10.1	Objectives	25
10.2	Prerequisites	25
10.3	General requirements	25
10.4	Architecture and development process design or refinement	26
10.5	Examples of architectural and development measures for AI systems	27
10.6	Work products	30
11	Data-related considerations	31
11.1	Objectives	31
11.2	Prerequisites and supporting information	31
11.3	General requirements	31
11.4	Dataset life cycle	32
11.4.1	Datasets and the AI safety lifecycle	32
11.4.2	Reference dataset lifecycle	33
11.4.3	Dataset safety analysis	34
11.4.4	Dataset requirements development	40
11.4.5	Dataset design	43
11.4.6	Dataset implementation	44
11.4.7	Dataset verification	45
11.4.8	Dataset validation	46
11.4.9	Dataset maintenance	46
11.5	Work products	47
12	Verification and validation of the AI system	47
12.1	Objectives	47
12.2	Prerequisites and supporting information	48
12.3	General requirements	48
12.4	AI/ML specific challenges to verification and validation	50
12.5	Verification and validation of the AI system	51
12.5.1	Scope of verification and validation of the AI system	51
12.5.2	AI component testing	1
12.5.3	Methods for testing the AI component	3
12.5.4	AI system integration and verification	5
12.5.5	Virtual testing vs physical testing	6
12.5.6	Evaluation of the safety-related performance of the AI system	7
12.5.7	AI system safety validation	8
12.6	Work products	8
13	Safety analysis of AI systems	9
13.1	Objectives	9
13.2	Prerequisites and supporting information	9
13.3	General requirements	10
13.4	Safety analysis of the AI system	10
13.4.1	Scope of the AI safety analysis	10
13.4.2	Safety analysis based on the results of testing	13

ISO/CD-PASDPAS 8800:2024-(en)

13.4.3 Safety analysis techniques.....	14
13.5 Work products.....	15
14 Measures during operation.....	15
14.1 Objectives.....	15
14.2 Prerequisites and supporting information.....	15
14.3 General requirements.....	16
14.4 Planning for operation and continuous assurance.....	16
14.4.1 Safety risk of the AI system during operation phase.....	16
14.4.2 Safety activities during the operation phase.....	17
14.5 Continual, periodic re-evaluation of the assurance argument.....	18
14.6 Measures to assure safety of the AI system during operation.....	19
14.6.1 General.....	19
14.6.2 Technical safety measures.....	19
14.6.3 Safe operation guidance and misuse prevention in the field.....	21
14.7 Field data collection.....	21
14.8 Evaluation and continuous development.....	22
14.8.1 Field risk evaluation.....	22
14.8.2 Countermeasures addressing field risk.....	23
14.8.3 AI re-training, re-validation, re-approval and re-deployment.....	24
14.9 Work products.....	25
15 Confidence in use of AI development frameworks and software tools used for AI model development.....	25
15.1 Objectives.....	25
15.2 Prerequisites and supporting information.....	25
15.3 General requirements.....	25
15.4 Confidence in the use of AI development frameworks.....	26
15.5 Confidence in the use of tools used to support the AI safety lifecycle.....	28
15.6 Principles for data-driven AI model training and evaluation.....	28
15.7 Work products.....	29
Annex A (informative) Overview and workflow of ISO PAS 8800.....	30
Annex B (informative) Example assurance argument structure for an AI-based vehicle function.....	35
B.1 General.....	35
B.2 Assurance argument pattern for supervised machine learning.....	35
B.3 Use of assurance claim points to increase confidence in the assurance argument.....	47
B.3.1 General remarks on the use of assurance claim points.....	47
B.3.2 Example assurance claim points to support assumptions or context: ACP-A2 for assumption A2.....	47
B.3.3 Example assurance claim point to support inference: ACP-S1 for strategy S1.....	48
B.3.4 Example assurance claim point to support evidence: ACP-E5.....	48
Annex C (informative) ISO 26262 gap analysis for ML.....	50
C.1 General.....	50
C.2 ISO 26262-4:2018 Tailoring and Guidance for ML.....	50
C.3 ISO 26262-6:2018 Tailoring for ML.....	51
Annex D (informative) Detailed considerations on safety-related properties of AI systems.....	57
Annex E (informative) STAMP/STPA example.....	60
E.1 Overview.....	60
E.2 STPA Example.....	60
E.2.1 STPA Step 1: Defining the purpose and scope of the analysis.....	60
E.2.2 STPA step 2: Modelling of the control structure.....	60
E.2.3 STPA step 3: Identification of unsafe control actions.....	61
E.2.4 STPA step 4: Identification of causal scenarios.....	61

E.2.5	Identifying safety measures to mitigate the safety-related issues	63
Annex F (informative)	Identification of software units within NN-based systems	65
Annex G (informative)	Architectural and development measures for AI systems	68
G.1	Examples of architectural and development measures for AI systems	68
G.1.1	Selection of architectural and development measures	68
G.1.2	Measures for architectural redundancy	68
G.2	Qualitative and quantitative analysis of AI architectures	71
G.2.1	The role of architecture analysis	71
G.2.2	Identifying software units within AI architectures	72
G.3	Data distributions and their impacts on AI models	73
G.3.1	Data distribution related issues	73
G.3.2	Out of distribution data and its mitigation	73
G.3.3	Distributional shift and its mitigation	74
G.4	Training safety measures	76
G.4.1	Hyperparameter tuning	76
G.4.2	Robust learning	76
G.4.3	Transfer learning	77
G.4.4	Confidence calibration and uncertainty quantification of AI models	77
G.4.5	Verifying feature selection	78
G.4.6	Monitoring multiple scores	78
G.4.7	Attention or saliency maps	78
G.4.8	Interpretable latent features	79
G.4.9	Augmentation of data	79
G.5	Monitoring and AI system modification	80
G.5.1	Dynamic environment monitoring	80
G.5.2	AI model modification	80
G.6	Alignment of intention	82
G.7	Considerations related to the target execution environment	82
G.7.1	Demonstrating AI safety within the target execution environment	82
G.7.2	Optimization of parameters and optimization of architectural entities of AI components	82
G.7.3	Knowledge distillation or the teacher-student model	82
G.7.4	Analysis for differences	83
Annex H (informative)	Typical performance metrics for machine learning	84
Bibliography		89
Foreword		xi
Introduction		xii
1	Scope	1
2	Normative references	1
3	Terms and definitions	2
3.1	General AI-related definitions	2
3.2	Data-related definitions	8
3.3	General safety-related definitions	9
3.4	Safety: Root cause-, error- and failure-related definitions	12
3.5	Miscellaneous definitions	13
4	Abbreviated terms	15
5	Requirements for conformity	16
5.1	Purpose	16
5.2	General requirements	16

6	AI within the context of road vehicles system safety engineering and basic concepts .....	17
6.1	Application of the ISO 26262 series for the development of AI systems .....	17
6.2	Interactions with encompassing system-level safety activities .....	18
6.3	Mapping of abstraction layers between the ISO 26262 series, ISO/IEC 22989 and this document .....	23
6.4	Example architecture for an AI system .....	27
6.5	Types of AI models .....	28
6.6	AI technologies of a ML model .....	29
6.7	Error concepts, fault models and causal models .....	32
6.7.1	Cause-and-effect chain .....	32
6.7.2	Root cause classes .....	33
6.7.3	Error classification based on the safety impact .....	35
7	AI safety management .....	35
7.1	Objectives .....	35
7.2	Prerequisites and supporting information .....	36
7.3	General requirements .....	36
7.4	Reference AI safety life cycle .....	39
7.5	Iterative development paradigms for AI systems .....	2
7.6	Work products .....	3
8	Assurance arguments for AI systems .....	3
8.1	Objectives .....	3
8.2	Prerequisites and supporting information .....	4
8.3	General requirements .....	4
8.4	AI system-specific considerations in assurance arguments .....	5
8.5	Structuring assurance arguments for AI systems .....	6
8.5.1	Context of the assurance argument .....	6
8.5.2	Categories of evidence .....	7
8.6	The role of quantitative targets and qualitative arguments .....	9
8.7	Evaluation of the assurance argument .....	10
8.8	Work products .....	11
9	Derivation of AI safety requirements .....	11
9.1	Objectives .....	11
9.2	Prerequisites and supporting information .....	11
9.3	General requirements .....	12
9.4	General workflow for deriving safety requirements .....	12
9.5	Deriving AI safety requirements on supervised machine learning .....	17
9.5.1	The need for refined AI safety requirements .....	17
9.5.2	Derivation of refined AI safety requirements to manage uncertainty .....	18
9.5.3	Refinement of the input space definition for AI safety lifecycle .....	21
9.5.4	Restricting the occurrence of AI output insufficiencies .....	22
9.5.5	Metrics, measurements and threshold design .....	26
9.5.6	Considerations for deriving safety requirements .....	27
9.6	Work products .....	28
10	Selection of AI technologies, architectural and development measures .....	28
10.1	Objectives .....	28
10.2	Prerequisites .....	28
10.3	General requirements .....	28
10.4	Architecture and development process design or refinement .....	29
10.5	Examples of architectural and development measures for AI systems .....	30
10.6	Work products .....	35
11	Data-related considerations .....	35
11.1	Objectives .....	35

11.2	Prerequisites and supporting information	35
11.3	General requirements	35
11.4	Dataset life cycle	36
11.4.1	Datasets and the AI safety lifecycle	36
11.4.2	Reference dataset lifecycle	39
11.4.3	Dataset safety analysis	41
11.4.4	Dataset requirements development	47
11.4.5	Dataset design	49
11.4.6	Dataset implementation	51
11.4.7	Dataset verification	51
11.4.8	Dataset validation	52
11.4.9	Dataset maintenance	53
11.5	Work products	54
12	Verification and validation of the AI system	54
12.1	Objectives	54
12.2	Prerequisites and supporting information	54
12.3	General requirements	55
12.4	AI/ML specific challenges to verification and validation	58
12.5	Verification and validation of the AI system	59
12.5.1	Scope of verification and validation of the AI system	59
12.5.2	AI component testing	1
12.5.3	Methods for testing the AI component	3
12.5.4	AI system integration and verification	6
12.5.5	Virtual testing vs physical testing	6
12.5.6	Evaluation of the safety-related performance of the AI system	7
12.5.7	AI system safety validation	8
12.6	Work products	9
13	Safety analysis of AI systems	9
13.1	Objectives	9
13.2	Prerequisites and supporting information	10
13.3	General requirements	10
13.4	Safety analysis of the AI system	11
13.4.1	Scope of the AI safety analysis	11
13.4.2	Safety analysis based on the results of testing	14
13.4.3	Safety analysis techniques	15
13.5	Work products	16
14	Measures during operation	16
14.1	Objectives	16
14.2	Prerequisites and supporting information	16
14.3	General requirements	17
14.4	Planning for operation and continuous assurance	17
14.4.1	Safety risk of the AI system during operation phase	17
14.4.2	Safety activities during the operation phase	18
14.5	Continual, periodic re-evaluation of the assurance argument	20
14.6	Measures to assure safety of the AI system during operation	20
14.6.1	General	20
14.6.2	Technical safety measures	21
14.6.3	Safe operation guidance and misuse prevention in the field	22
14.7	Field data collection	22
14.8	Evaluation and continuous development	24
14.8.1	Field risk evaluation	24
14.8.2	Countermeasures addressing field risk	25
14.8.3	AI re-training, re-validation, re-approval and re-deployment	25



14.9	Work products .....	26
15	Confidence in use of AI development frameworks and software tools used for AI model development .....	26
15.1	Objectives .....	26
15.2	Prerequisites and supporting information .....	27
15.3	General requirements .....	27
15.4	Confidence in the use of AI development frameworks .....	27
15.5	Confidence in the use of tools used to support the AI-safety lifecycle .....	30
15.6	Principles for data-driven AI model training and evaluation .....	30
15.7	Work products .....	31
Annex A	(informative) Overview and workflow of this document .....	32
Annex B	(informative) Example assurance argument structure for an AI-based vehicle function .....	37
B.1	General .....	37
B.2	Assurance argument pattern for supervised machine learning .....	37
B.3	Use of assurance claim points to increase confidence in the assurance argument .....	56
B.3.1	General remarks on the use of assurance claim points .....	56
B.3.2	Example assurance claim points to support assumptions or context: ACP-A2 for assumption A2 .....	56
B.3.3	Example assurance claim point to support inference: ACP-S1 for strategy S1 .....	59
B.3.4	Example assurance claim point to support evidence: ACP-E5 .....	59
Annex C	(informative) ISO 26262 gap analysis for ML .....	62
C.1	General .....	62
C.2	ISO 26262-4:2018 Tailoring and Guidance for ML .....	62
C.3	ISO 26262-6:2018 Tailoring for ML .....	63
Annex D	(informative) Detailed considerations on safety-related properties of AI systems .....	69
Annex E	(informative) STAMP/STPA example .....	72
E.1	Overview .....	72
E.2	STPA Example .....	72
E.2.1	STPA Step 1: Defining the purpose and scope of the analysis .....	72
E.2.2	STPA step 2: Modelling of the control structure .....	72
E.2.3	STPA step 3: Identification of unsafe control actions .....	73
E.2.4	STPA step 4: Identification of causal scenarios .....	74
E.2.5	Identifying safety measures to mitigate the safety-related issues .....	75
Annex F	(informative) Identification of software units within NN-based systems .....	77
Annex G	(informative) Architectural and development measures for AI systems .....	80
G.1	Examples of architectural and development measures for AI systems .....	80
G.1.1	Selection of architectural and development measures .....	80
G.1.2	Measures for architectural redundancy .....	80
G.2	Qualitative and quantitative analysis of AI architectures .....	84
G.2.1	The role of architecture analysis .....	84
G.2.2	Identifying software units within AI architectures .....	85
G.3	Data distributions and their impacts on AI models .....	85
G.3.1	Data distribution related issues .....	85
G.3.2	Out of distribution data and its mitigation .....	85
G.3.3	Distributional shift and its mitigation .....	86
G.4	Training safety measures .....	88
G.4.1	Hyperparameter tuning .....	88
G.4.2	Robust learning .....	89
G.4.3	Transfer learning .....	89
G.4.4	Confidence calibration and uncertainty quantification of AI models .....	90
G.4.5	Verifying feature selection .....	90

ISO/~~CD-PASDPAS~~ 8800:2024(en)

<u>G.4.6 Monitoring multiple scores.....</u>	<u>91</u>
<u>G.4.7 Attention or saliency maps.....</u>	<u>91</u>
<u>G.4.8 Interpretable latent features.....</u>	<u>91</u>
<u>G.4.9 Augmentation of data.....</u>	<u>91</u>
<u>G.5 Monitoring and AI system modification.....</u>	<u>93</u>
<u>G.5.1 Dynamic environment monitoring.....</u>	<u>93</u>
<u>G.5.2 AI model modification.....</u>	<u>93</u>
<u>G.6 Alignment of intention.....</u>	<u>94</u>
<u>G.7 Considerations related to the target execution environment.....</u>	<u>95</u>
<u>G.7.1 Demonstrating AI safety within the target execution environment.....</u>	<u>95</u>
<u>G.7.2 Optimization of parameters and optimization of architectural entities of AI components.....</u>	<u>95</u>
<u>G.7.3 Knowledge distillation or the teacher-student model.....</u>	<u>95</u>
<u>G.7.4 Analysis for differences.....</u>	<u>95</u>
<u>Annex H (informative) Typical performance metrics for machine learning.....</u>	<u>96</u>
<u>Bibliography.....</u>	<u>102</u>

iTeh Standards  
(<https://standards.iteh.ai>)  
Document Preview

ISO/DPAS 8800

<https://standards.iteh.ai/catalog/standards/iso/f16a69d8-8442-461c-a36b-5756f4e356a7/iso-dpas-8800>

## 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](http://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](http://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](http://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](http://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—<sup>1)</sup>, 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 <a href="#">Clause 5</a> for details.
ISO 21448	This document is a tailoring or extension of ISO 21448 for AI elements of the system. See <a href="#">Clause 5</a> for details.

<sup>1)</sup> Under preparation. Stage at the time of publication: ISO/DTS 5083.

ISO/~~CD-PASDPAS~~ 8800:~~2024~~:(en)

Publication	Relationship with this document
ISO TS 5083: <del>2024</del> <del>20</del>	ISO TS 5083: <del>20</del> , 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.

ITeh Standards  
(<https://standards.iteh.ai>)  
Document Preview

[ISO/DPAS 8800](https://standards.iteh.ai/catalog/standards/iso/fl6a69d8-8442-461c-a36b-5756f4e356a7/iso-dpas-8800)

<https://standards.iteh.ai/catalog/standards/iso/fl6a69d8-8442-461c-a36b-5756f4e356a7/iso-dpas-8800>