

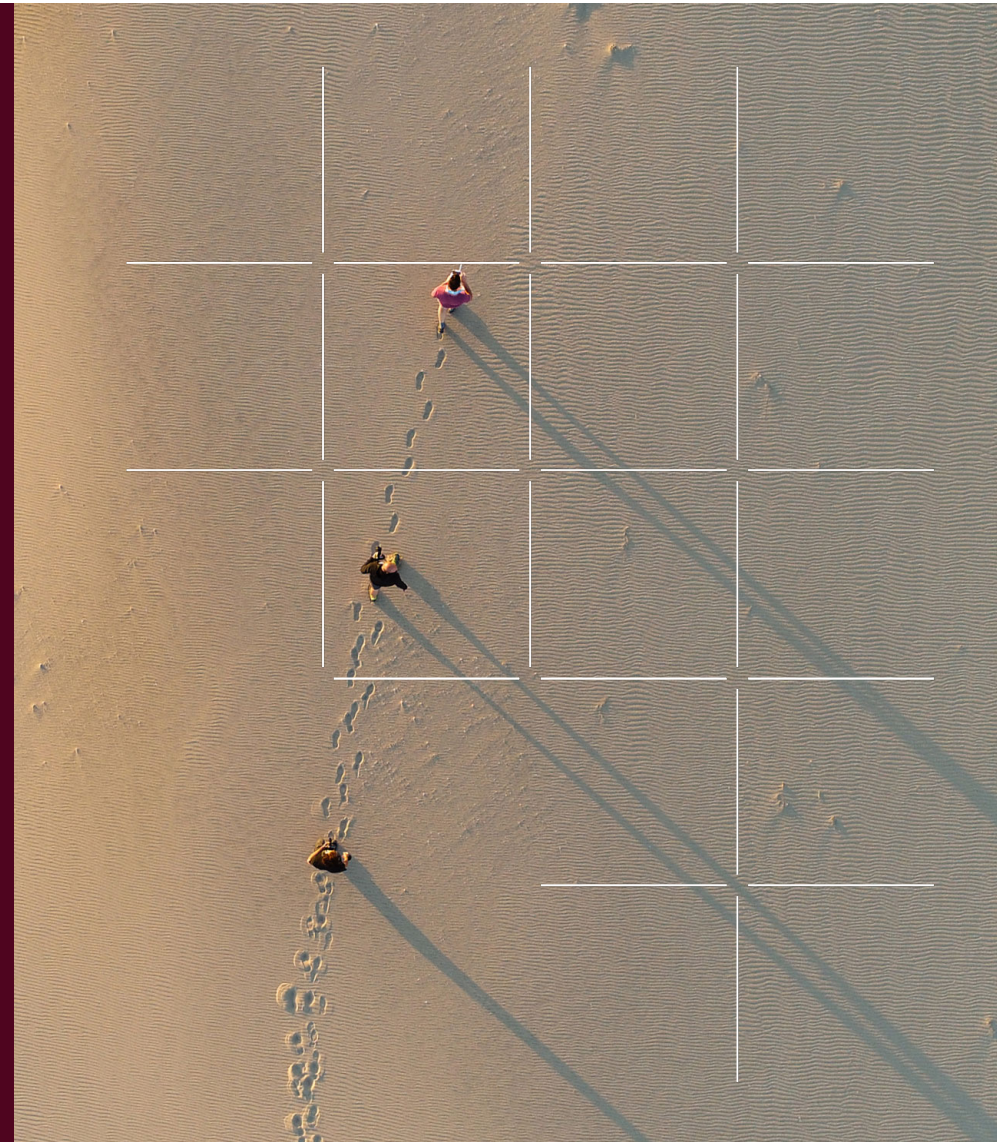


ISO TC22/SC32/WG14, ISO PAS 8800 Road Vehicles – Safety and AI

Prof. Simon Burton

Convenor, ISO TC22/SC32/WG14
Chair of Systems Safety, University of York

2025-10, Online



Automotive safety and AI

All AI/ML models fail, sometimes, and under certain conditions:

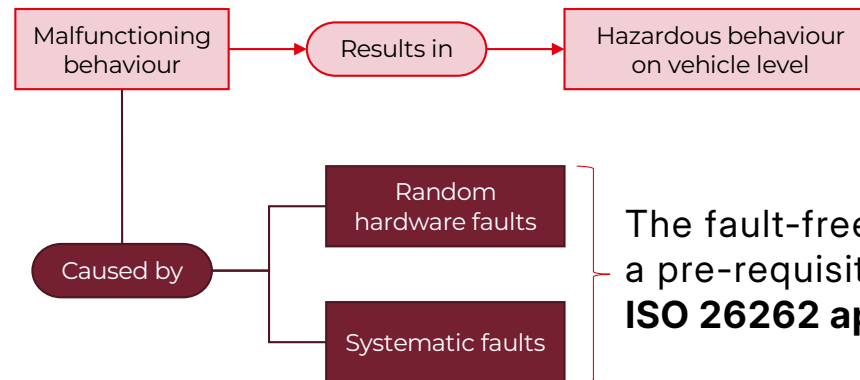
- How can we nevertheless build systems that fully exploit the **innovative potential of AI**
- ...whilst remaining **safe** and **trusted** by both the general public and regulators?



Applicability of existing safety standards

ISO 26262: Functional safety

*"Absence of unreasonable risk due to hazards caused by **malfunctioning behaviour** of the electrical and/or electronic systems"*



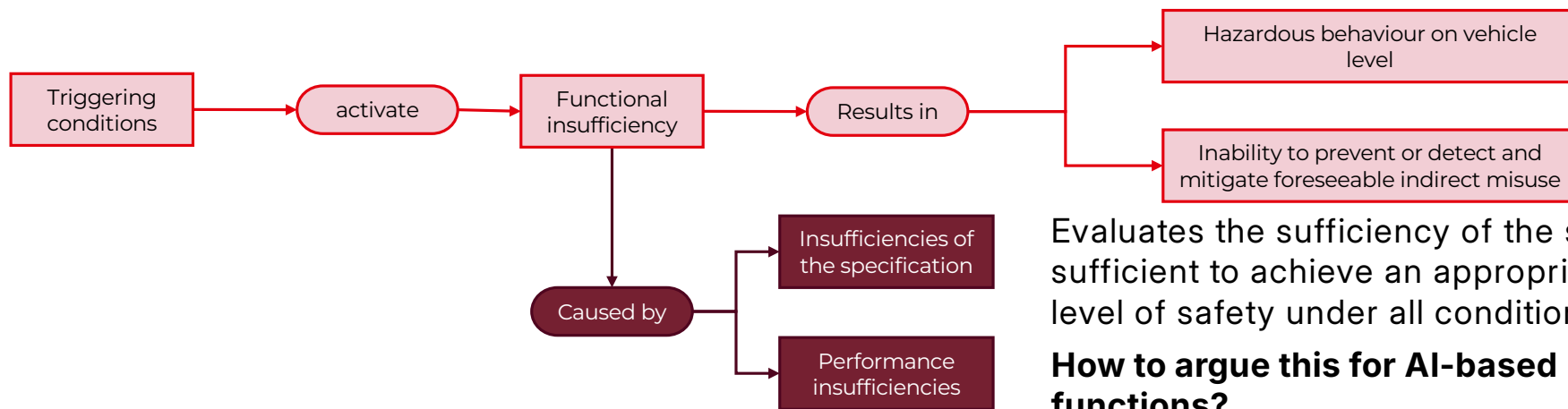
The fault-free execution of the AI models is a pre-requisite to AI safety
ISO 26262 applies to AI-based systems



Applicability of existing safety standards

ISO 21448: Safety of the intended functionality (SOTIF)

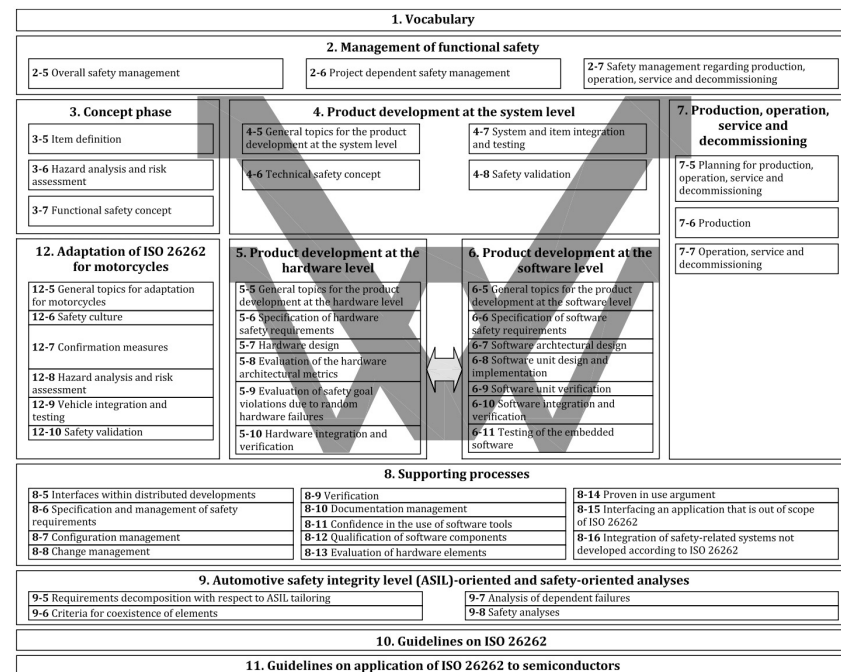
"Absence of unreasonable risk due to hazards resulting from functional insufficiencies of the intended functionality or by reasonably foreseeable misuse by road users"



ISO PAS 8800 - Motivation

"Classic" safety-critical software: Engineered to be safe by design...

- for comparatively well-understood and well-defined tasks...
- Evidence from the specification and design of the software supported by...
- White-box testing, review and code analysis and
- Black-box testing based on assumptions about equivalence classes of errors
- Validation in a relatively stable environment...



ISO 26262 safety life cycle



ISO PAS 8800 - Motivation

"Classic" safety-critical software:
Engineered to be safe by design...

- ... see existing standards...

1. Vocabulary		
2. Management of functional safety		
2-5 Overall safety management	2-6 Project dependent safety management	2-7 Safety management regarding production, operation, service and decommissioning
3. Concept phase	4. Product development at the system level	7. Production, operation, service and decommissioning
3-5 Item definition	4-5 General topics for the product development at the system level	4-7 System and item integration and testing
3-6 Hazard analysis and risk assessment	4-6 Technical safety concept	4-8 Safety validation
3-7 Functional safety concept		7-5 Planning for production, operation, service and decommissioning
		7-6 Production
		7-7 Operation, service and decommissioning
12. Adaptation of ISO 26262 for motorcycles	5. Product development at the hardware level	6. Product development at the software level
12-3 General topics for adaptation for motorcycles	5-5 General topics for the product development at the hardware level	6-5 General topics for the product development at the software level
12-6 Safety culture	5-6 Specification of hardware safety requirements	6-6 Specification of software safety requirements
12-7 Confirmation measures	5-7 Hardware design	6-7 Software architectural design
12-8 Hazard analysis and risk assessment	5-8 Evaluation of the hardware architectural metrics	6-8 Software unit design and implementation
12-9 Vehicle integration and testing	5-9 Evaluation of safety goal violations due to hardware failures	6-9 Software unit verification
12-10 Safety validation	5-10 Hardware integration and verification	6-10 Software integration and verification
		6-11 Testing of the embedded software
8. Supporting processes		
8-5 Interfaces within distributed developments	8-9 Verification	8-14 Proven in use argument
8-6 Specification and management of safety requirements	8-10 Documentation management	8-15 Interfacing an application that is out of scope of ISO 26262
8-7 Configuration management	8-11 Confidence in the use of software tools	8-16 Integration of safety-related systems not developed according to ISO 26262
8-8 Change management	8-12 Qualification of software components	
	8-13 Evaluation of hardware elements	
9. Automotive safety integrity level (ASIL)-oriented and safety-oriented analyses		
9-5 Requirements decomposition with respect to ASIL tailoring	9-7 Analysis of dependent failures	
9-6 Criteria for coexistence of elements	9-8 Safety analyses	
10. Guidelines on ISO 26262		
11. Guidelines on application of ISO 26262 to semiconductors		

AI-based systems...

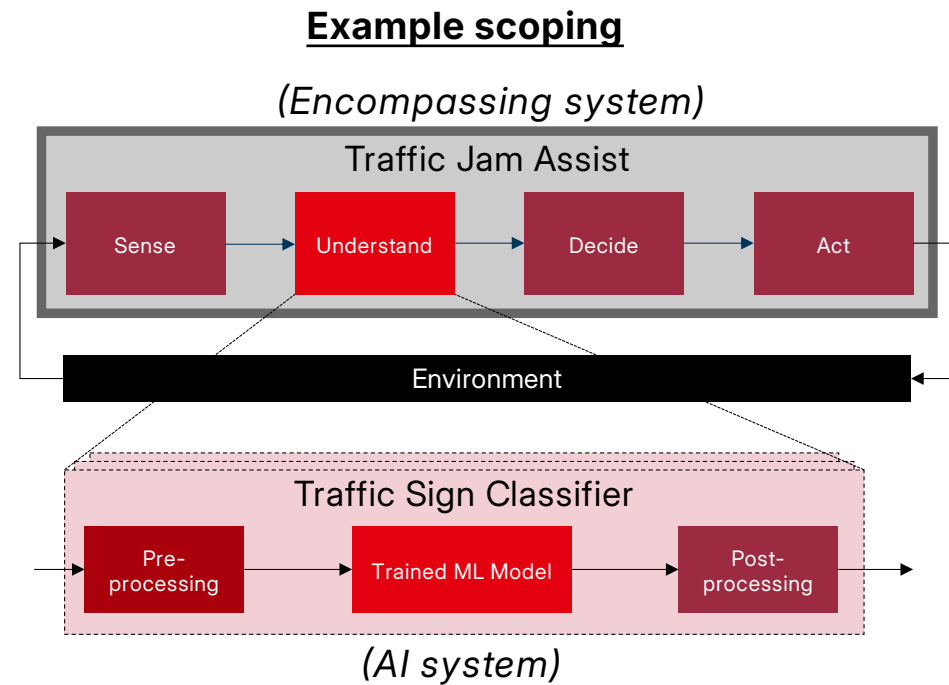
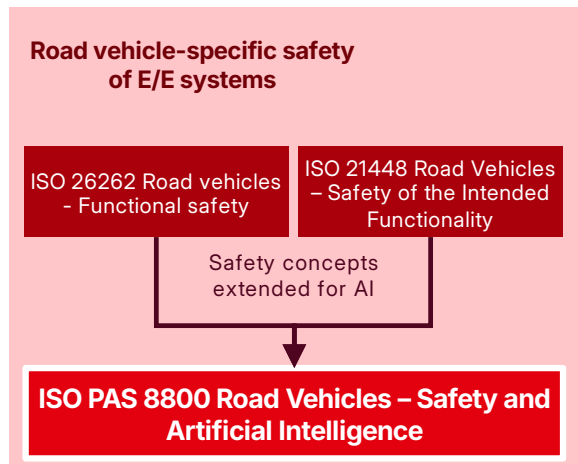
- Reliance on quantitative test evidence and ML metrics based on statistical assumptions about error distribution and the operating conditions
- Is this enough?

Specific guidance needed on how to achieve **an equivalent level of Safety Integrity** to existing safety-critical software components **for AI/ML-based systems**



ISO PAS 8800 Road vehicles – Safety and AI

Overview



Scope of ISO TC22/SC32/WG 14

Road vehicles – Safety and Artificial Intelligence

Extension of concepts from ISO 26262 (functional safety) and ISO 21448 (safety of the intended functionality)

- Applicable to a **range of applications** (not only automated driving)
- Applicable to a **range of AI/ML technologies** (not just DNNs for object detection)
- Contains **normative requirements** for phases of the AI Safety Lifecycle
- **Informative guidance and examples** related to methods and metrics to support the lifecycle

Continuous assurance during operation

Encompassing AI system:

Pre- and postprocessing to reduce impact of AI errors, consideration of known insufficiencies in system requirements, assurance argument

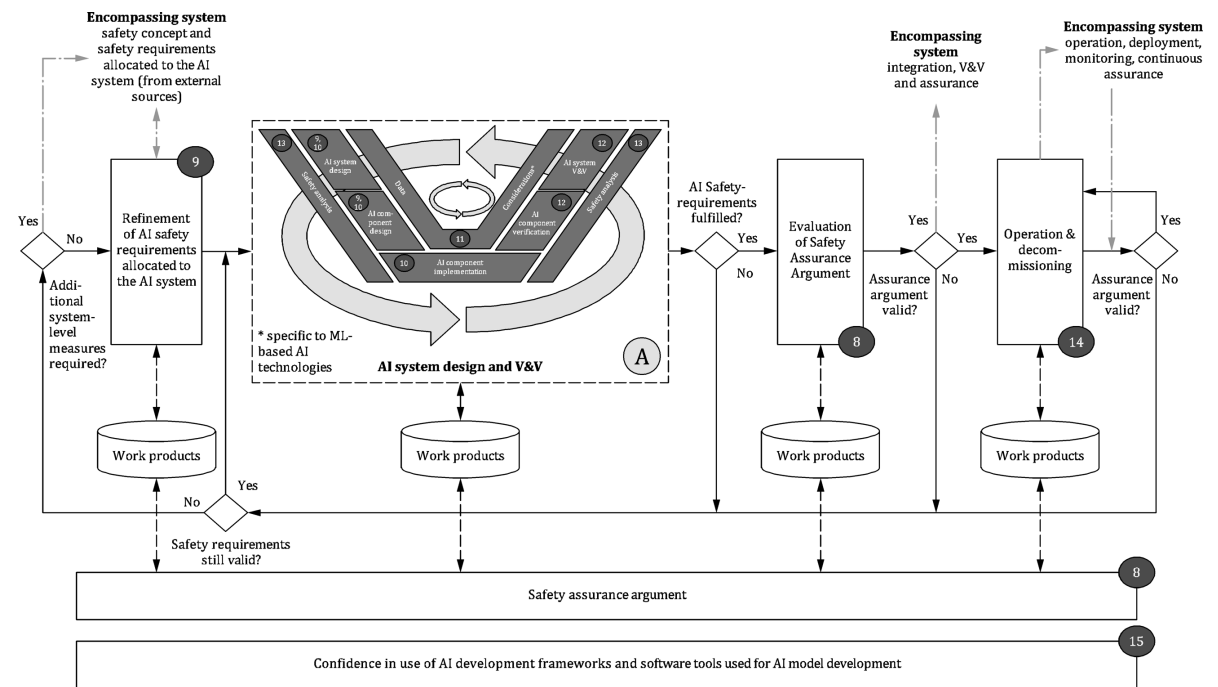
AI model:

Specification of safety-related (quantitative) properties, measures to reduce technical uncertainty, V&V, Safety Analysis



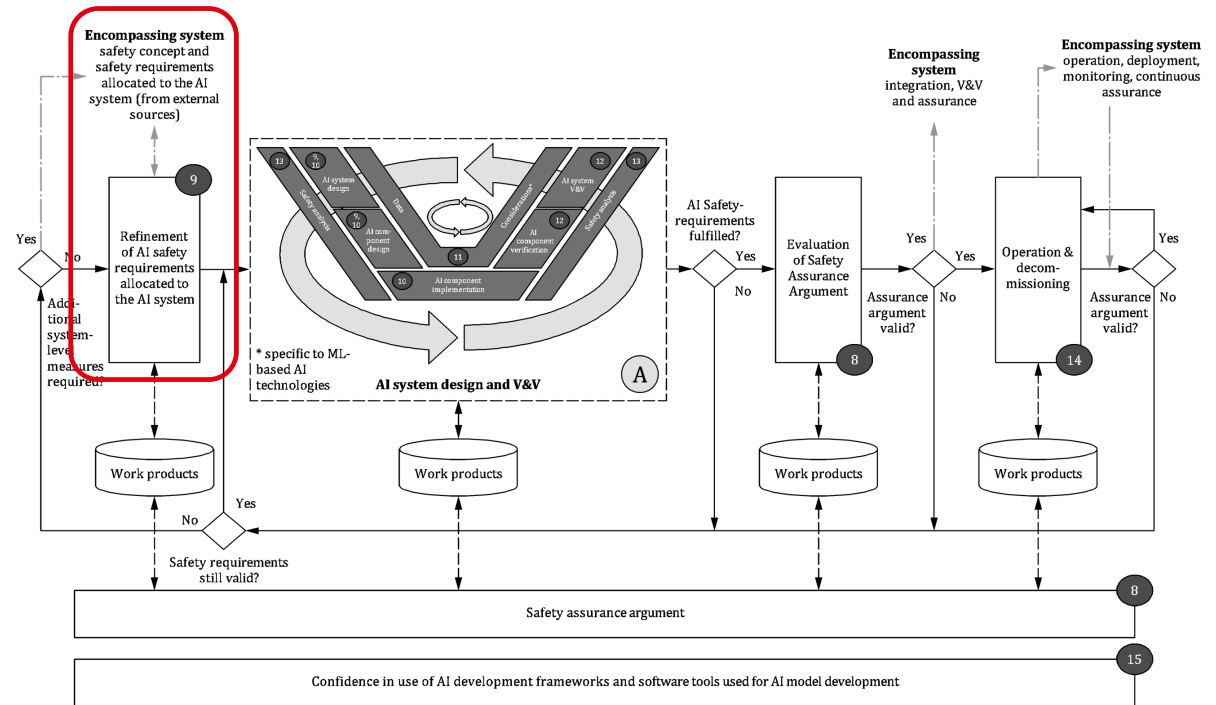
Highlights of ISO PAS 8800

- Definition of an **iterative AI safety lifecycle** summarising safety-relevant activities, including continuous safety assurance during **operation**



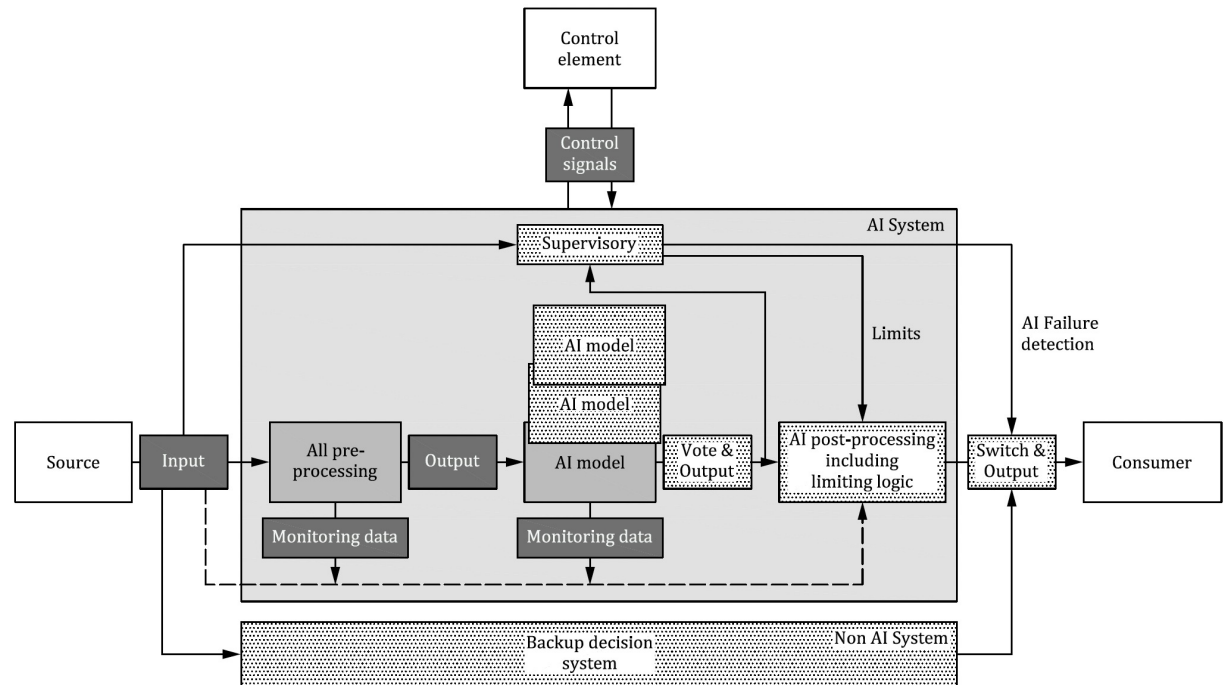
Highlights of ISO PAS 8800

- Definition of a **fault model** and **safety-related properties** used to define detailed safety requirements of the AI systems (compatible with ISO 26262 and ISO 21448)
- **Known limitations** of the AI system are used to design mitigating measures at the encompassing system level







Highlights of ISO PAS 8800

- Definition of **development** and **architectural measures** for achieving the safety-related properties of AI systems

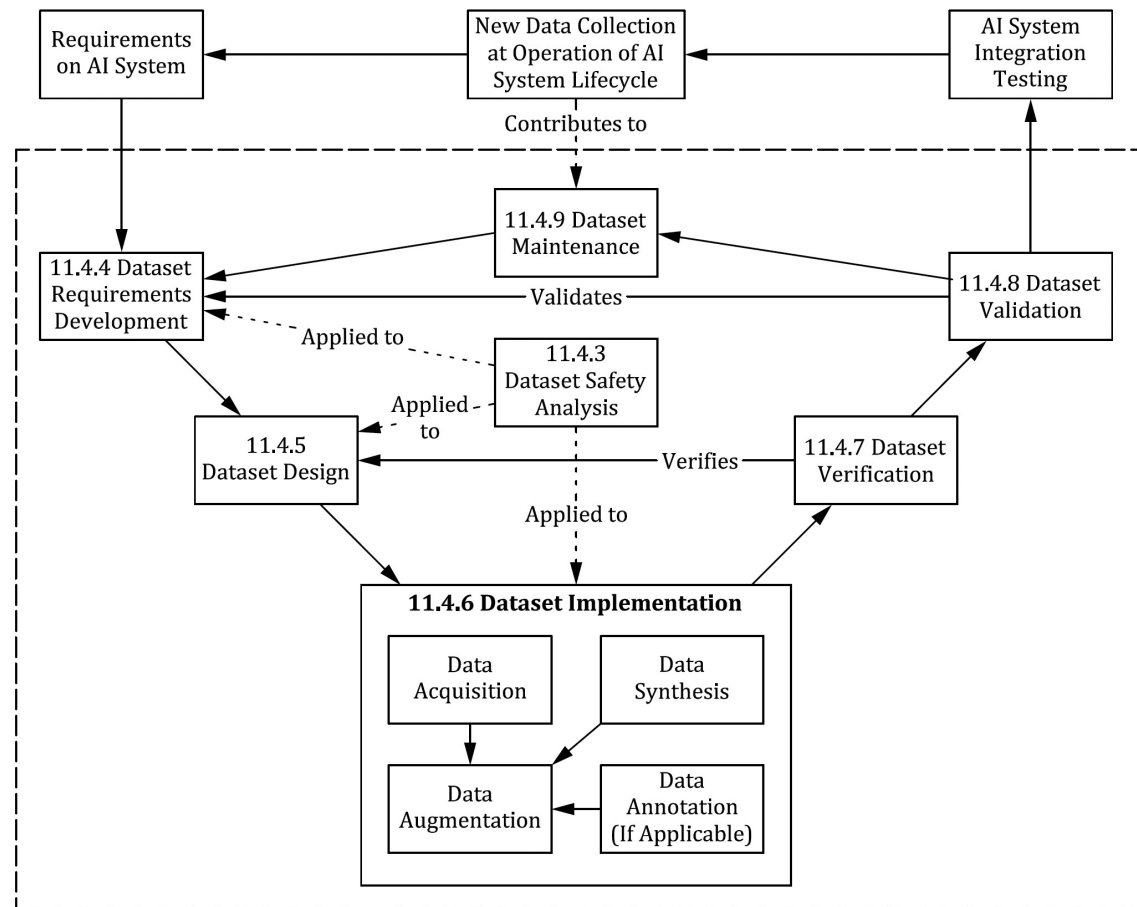


Key

-  architecture redundancy elements
-  AI system
-  AI component
-  element not belonging to this AI system

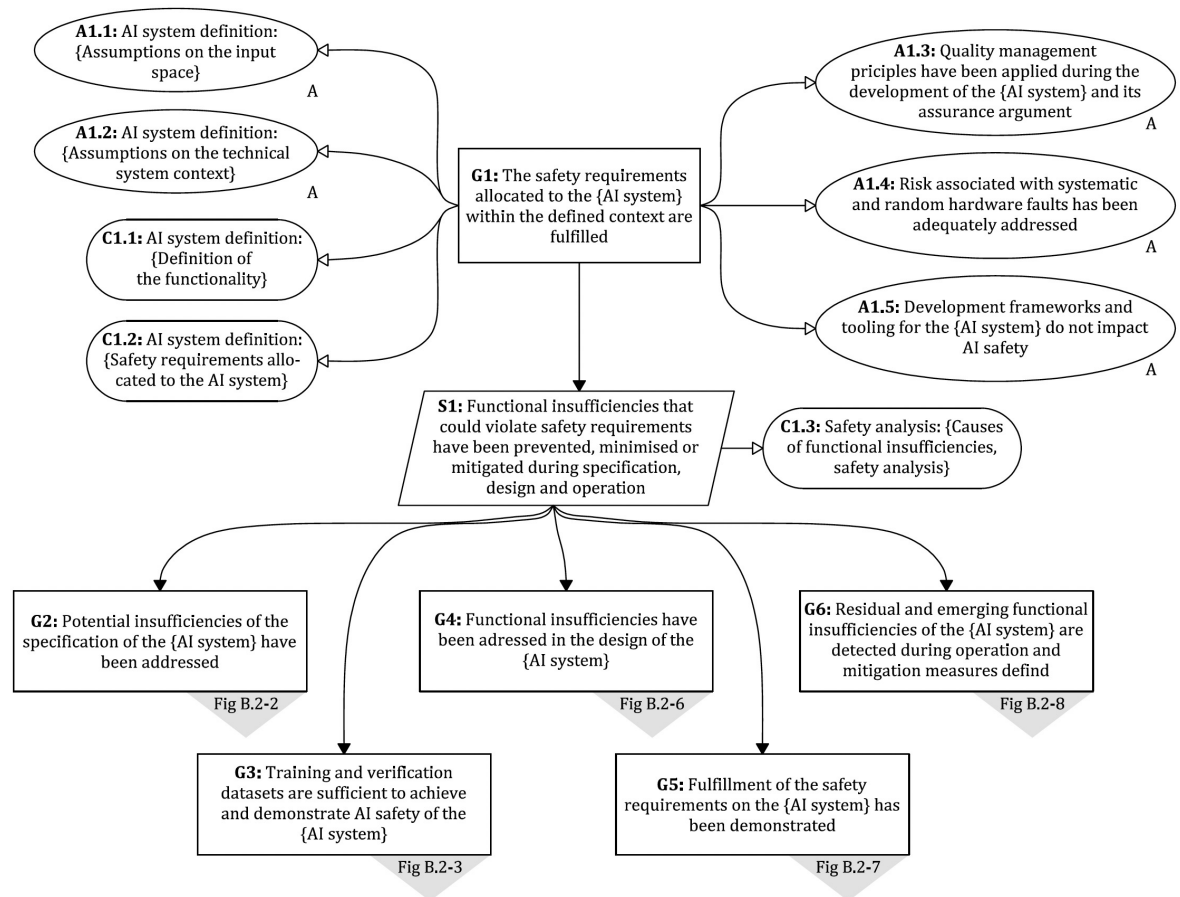
Highlights of ISO PAS 8800

- **Safety-related properties of data** and data management lifecycle

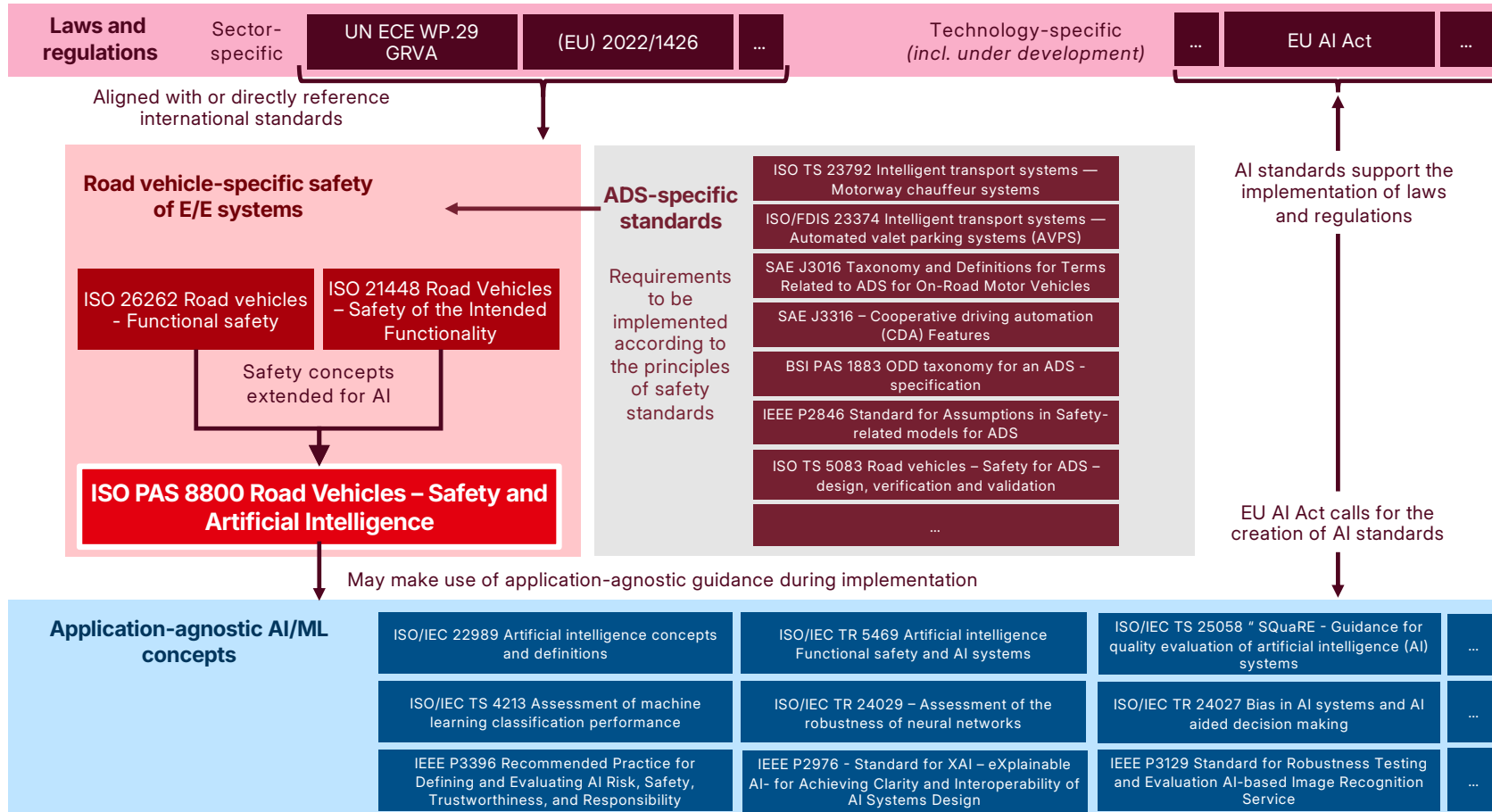


Highlights of ISO PAS 8800

- **Safety analyses and structured assurance arguments** for justifying acceptable residual risk associated with the AI system, including **continual assurance during operation**



A complex, evolving standards landscape



Current status of ISO PAS 8800

Approved 2024-10-16, released on 2024-12-12

Many thanks to > 130 experts from 13 different countries for their contributions

Result of voting
<p>P-Members voting: 20 in favour out of 20 = 100 % (requirement \geq 50%) <i>(P-Members having abstained are not counted in this vote.)</i></p>
<p>Member bodies voting: 0 negative votes out of 20 = 0 % (requirement \leq 25%)</p>
Approved

Contents		
Foreword		
Introduction		
1 Scope		
2 Normative references		
3 Terms and definitions		
3.1 General AI-related definitions		
3.2 Data-related definitions		
3.3 General safety-related definitions		
3.4 Safety: Root cause-, error-and failure-related		
3.5 Miscellaneous definitions		
4 Abbreviated terms		
5 Requirements for conformity		
5.1 Purpose		
5.2 General requirements		
6 AI within the context of road vehicles system safe		
6.1 Application of the ISO 26262 series for the dev		
6.2 Interactions with encompassing system-level		
6.3 Mapping of abstraction layers between the IS document		
6.4 Example architecture for an AI system		
6.5 Types of AI models		
6.6 AI technologies of a ML model		
6.7 Error concepts, fault models and causal model		
6.7.1 Cause-and-effect chain		
6.7.2 Root cause classes		
6.7.3 Error classification based on the safety		
7 AI safety management	28	
7.1 Objectives	28	
7.2 Prerequisites and supporting information	28	
7.3 General requirements	28	
7.4 Reference AI safety life cycle	31	
7.5 Iterative development paradigms for AI systems	33	
7.6 Work products	34	
8 Assurance arguments for AI systems	35	

ISO/PAS 8800

Road vehicles — Safety and artificial intelligence

Véhicules routiers — Sécurité et intelligence artificielle

First edition 2024-12



ISO TC22/SC32/WG14 – Current phase of work

- Scoping of 2nd Edition of ISO 8800 as full international
- Scoping of TR with application examples for driver assistance and automated driving perception functions
- Scoping of TR with more specific methods and metrics guidance
- Liaisons, in particular with CEN/CENELEC JTC21 regarding implementation of the EU AI Act standardization requests (see next slides)
- Other topics: E2E learning, data considerations,...




Is ISO PAS 8800 compliance sufficient to fulfill the EU AI act?

EU AI Act requirements on high-risk systems:

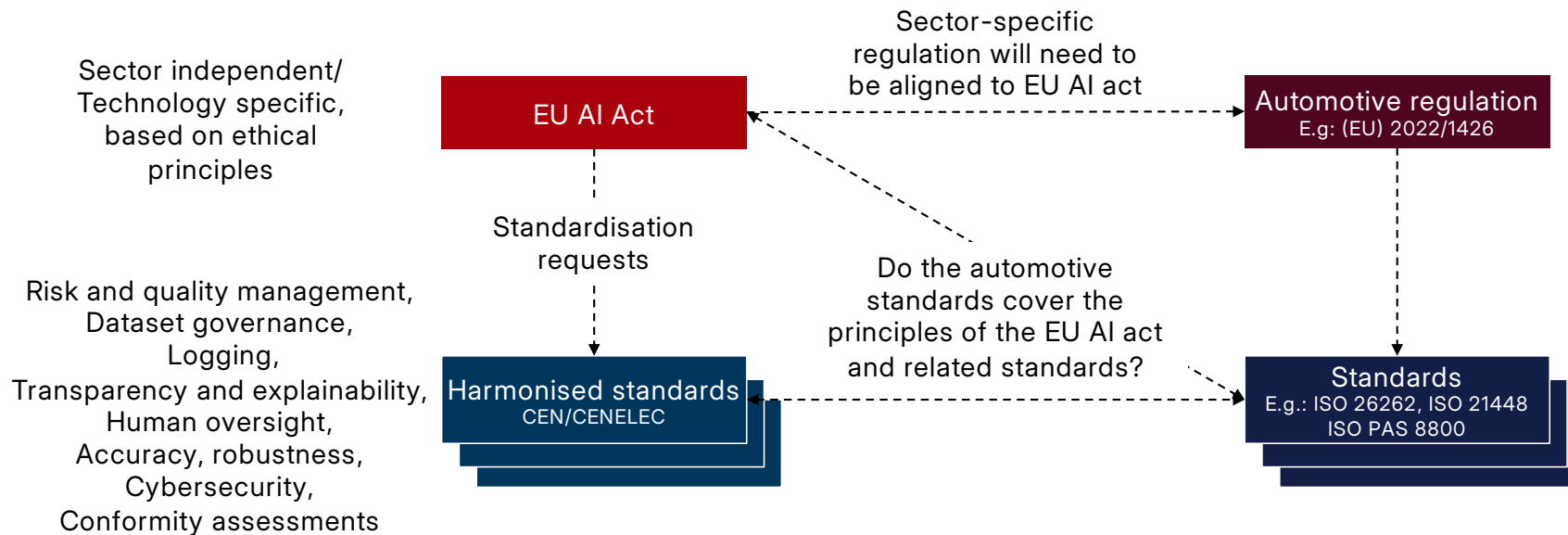
- Article 9 - Risk management system
- Article 10 - Data and data governance
- Article 11 – Technical documentation
- Article 12 – Record keeping
- Article 13 - Transparency and provision of information to users
- Article 14 - Human oversight
- Article 15 - Accuracy, robustness and cybersecurity



<small>Single user licence only, copying and networking prohibited</small>	
	Publicly Available Specification
	ISO/PAS 8800
Road vehicles — Safety and artificial intelligence <small>Véhicules routiers — Sécurité et intelligence artificielle</small>	First edition 2024-12



Is ISO PAS 8800 compliance sufficient to fulfill the EU AI act?



Is ISO PAS 8800 compliance sufficient to fulfill the EU AI act?

- EU AI act and ISO PAS 8800 both take a risk-based approach
- Many requirements are similar e.g. on data management and accuracy, robustness
- **However:**
 - Automotive safety standards derive safety requirements from **system-level functional risks** and then refines based on engineering decisions (**Vertical standards**)
 - EU AI Act assumes the **use of a specific technology** and defines relevant properties accordingly (**Horizontal standards**) ...but many requirements in EU AI Act appear motivated by the usage (application) and not the technology itself
- **→ Many similarities but also subtle differences that require deeper understanding and interpretation**



Next steps

- Formal liaison established between ISO TC22/SC32/WG14 and CEN/CENELEC JTC21 to allow for coordination in the implementation of the EU AI Act standardization requests
- Task force established within ISO TC22/SC32/WG14 to coordinate the liaison activities
- **A coordinated and clearly expressed automotive position on the EU AI act (and comparable regional regulations) w.r.t. safety-relevant functions is required**



Road vehicles Safety and AI – The big picture

Summary

- ISO PAS 8800 is an important first step in providing pragmatic guidance for the **integration of AI/ML components into safety-relevant in-vehicle systems**
- Focus is on the **safety management system** and **AI safety lifecycle** with normative requirements on each phase
- Must be seen (and used) within the wider context of ISO 26262 and ISO 21448 and the wider evolving **standardisation and regulatory landscape**
 - E.g. Application-specific standards for deriving application-specific safety requirements on AI systems
 - E.g. Sector-independent standards for selecting appropriate design and verification methods and metrics





Thank you.

Making lives *easier*, *safer* and *better*.

Prof. Simon Burton:
simon.burton@york.ac.uk
Convenor ISO TC22/SC 32/WG 14

[iso.org](https://www.iso.org)