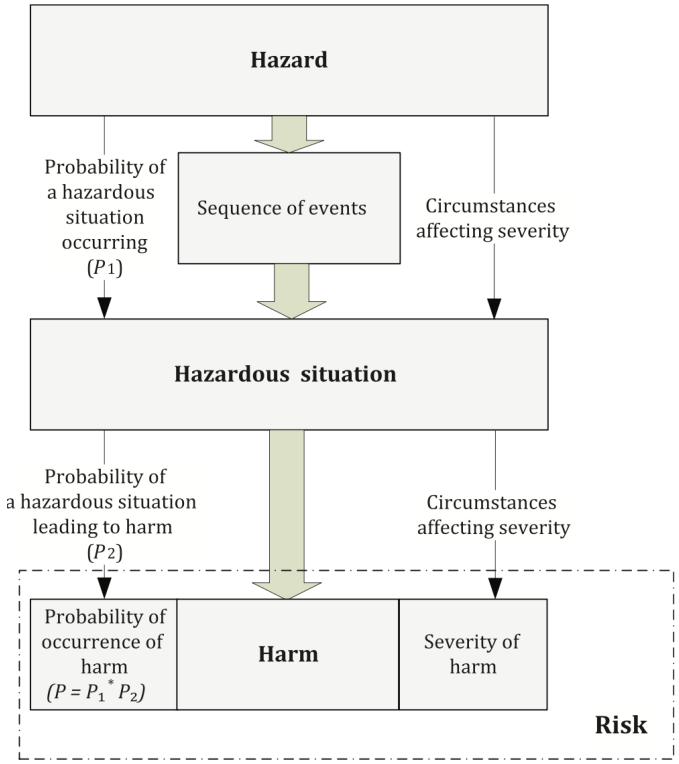
ISO 14971


Hazard
Probability of a hazardous situation occurring (P_1)
Sequence of events
Circumstances affecting severity
Hazardous situation
Probability of a hazardous situation leading to harm (P_2)
Circumstances affecting severity
Probability of occurrence of harm ($P = P_1 * P_2$)
Harm
Severity of harm
Risk

ISO 14971

Table C.3 — Relationship between *hazards*, foreseeable sequences of events, *hazardous situations* and the *harm* that can occur

Hazard	Foreseeable sequence of events	*Hazardous situation*	*Harm*
Electromagnetic energy (high voltage)	(1) Electrode cable unintentionally plugged into power line receptacle	Line voltage appears on electrodes	Serious burns Heart fibrillation
Chemical (volatile solvent, embolus)	(1) Incomplete removal of volatile solvent used in manufacturing (2) Solvent residue converts to gas at body temperature	Development of gas embolism (bubbles in the blood stream) during dialysis	Infarct Brain damage
Biological (microbial contamination)	(1) Inadequate instructions provided for decontaminating re-used anaesthesia tubing (2) Contaminated tubing used during anaesthesia	Bacteria released into airway of patient during anaesthesia	Bacterial infection
Functionality (no delivery)	(1) Electrostatically charged patient touches infusion pump (2) Electrostatic discharge (ESD) causes pump and pump alarms to fail	Failure to deliver insulin to patient with elevated blood glucose level, no warning given	Minor organ damage Decreased consciousness
Functionality (no output)	(1) Implantable defibrillator battery reaches the end of its useful life (2) Inappropriately long interval between clinical follow-up visits	Defibrillator cannot deliver shock when an arrhythmia occurs	Death
Measurement (incorrect information)	(1) Measurement error (2) No detection by user	Incorrect information reported to clinician, leading to misdiagnosis and/or lack of proper therapy	Progression of disease Serious injury

ISO 23894

Table C.1 — Risk management and AI system life cycle

→ Risk management AI system Life cycle ↓	AI risk management framework (Clause 5)	AI risk management process (Clause 6)				
		Scope, context and criteria	**Risk assessment**	**Risk treatment**	**Monitoring and review**	**Recording and reporting**
Organizational level activities related to risk management	Governing body sets directions for AI risk management. Top management commits. High-level risk management appetite and general criteria are established.	Feedback reports from AI systems' risk management processes are being received and processed. As a result, the organizational risk management framework is being improved by extending and refining of the organization's risk management tools:				
		A catalogue of risk criteria.	A catalogue of potential risk sources. A catalogue of techniques for risk sources' assessment and measurements.	A catalogue of known or implemented mitigation measures.	A catalogue of known or implemented techniques for monitoring and controlling AI systems.	A catalogue of established methods and defined formats for tracing, recording, reporting, and sharing the information about AI systems with internal and external stakeholders.
Inception	Governing body examines the AI system objectives in the context of the organization's and the stakeholders' principles and values, Based on a (typically multi-layer) analysis, determines whether the AI system is feasible and addresses the problem the organization seeks to solve.	The AI system risk management process and the system's risk criteria are established through customization of the organization's risk management framework.	Risk sources specific to the AI system are identified (potentially in a multi-layered manner) and described in detail.	A detailed risk treatment plan is established. Potentially, "proof of concept" methods are defined.	Necessary "proof of concept" methods are implemented, tested and evaluated.	The analysis with its results and the recommendation are recorded and communicated to the top management.
Design and development	Governing body continually re-assesses the objectives, the efficacy and the feasibility of the system based on received feedback reports.	Potentially, the AI system risk criteria is modified as a result of the feedback reports.	The risk assessment is performed continuously (potentially on multiple layers).	The risk treatment plan is implemented. The risk treatment and the (remaining) risks assessment continue until the established risk criteria are met.	During the testing, verification and validation the risk treatment plan for the system's components as well as for the whole system is assessed and adjusted.	The results are recorded and fed back to the relevant risk management process activities. As necessary, the conclusions are communicated to the management chain and to the governance body.
Verification and validation						

ISO 23894

Table C.1 *(continued)*

→ Risk management AI system Life cycle ↓	AI risk management framework (Clause 5)	AI risk management process (Clause 6)				
		Scope, context and criteria	**Risk assessment**	**Risk treatment**	**Monitoring and review**	**Recording and reporting**
Deployment	Governing body continually re-assesses the objectives and the feasibility of the system based on received feedback reports.	The AI system risk criteria and the risk management process are adjusted for the necessary "configuration" changes.	The risk assessment is performed continuously (potentially on multiple layers).	The risk treatment plan is potentially updated due to "configuration" changes and implemented. The risk treatment and the (remaining) risks assessment continue until the established risk criteria are met.	The AI system's risk treatment plan is being re-assessed to allow for necessary adjustments.	
Operation, monitoring **Continuous validation**	Governing body continually re-assesses the objectives and the feasibility of the system based on received feedback reports.	Potentially, the AI system risk criteria is modified as a result of the feedback reports.	The system's risk assessment plan is potentially adjusted for risk criteria changes.	The system's risk treatment plan is potentially adjusted for risk changes in risk assessment outcomes.	The risk treatment plan for the system's components is assessed and adjusted.	
Re-evaluation	Governing body re-examines the AI system objectives and their relation to the organization's and the stakeholders' principles and values, Based on the analysis, determines whether the AI system is feasible.	The AI system risk management process and the system's risk criteria are re-evaluated against any potential changes to the specific purpose and scope of the AI system, outcome of operation monitoring and new regulatory requirements	The list of existing risk sources specific to the AI system are examined for relevance and any possible gaps.	The risk treatment plan is potentially updated. The risk treatment and the (remaining) risks assessments continue until the established risk criteria are met.	The AI system's risk treatment plan is being re-assessed to allow for necessary adjustments.	
Retirement or replacement Triggers a new risk management process with new objectives, risks and their mitigation.	Governing body re-examines the AI system objectives based on the analysis, determines whether the AI system retirement or replacement is feasible.	The AI system risk management retirement process and the system's retirement risk criteria are established.	Risk sources specific to the AI system retirement are identified and described in detail.	Detailed risk treatment plan is established.	Necessary "proof of concept" methods are implemented, tested and evaluated.	

ISO 26262

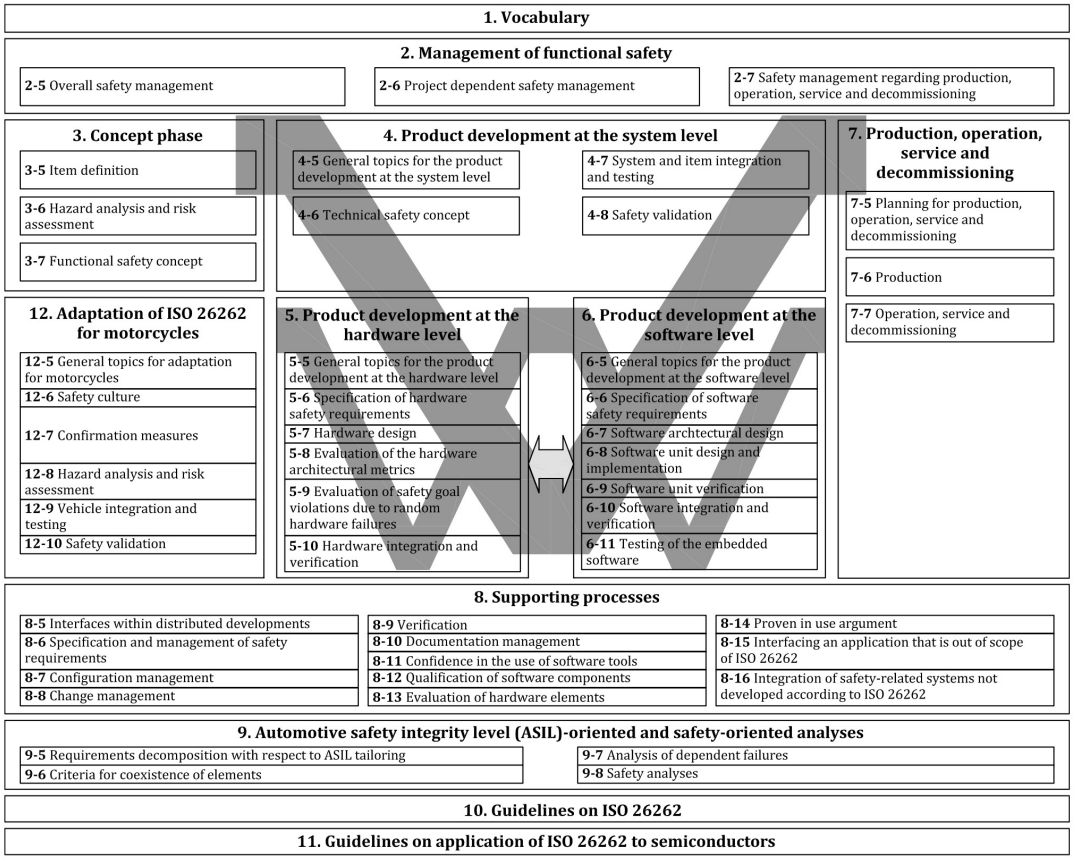


Figure 1 — Overview of the ISO 26262 series of standards

ISO 26262

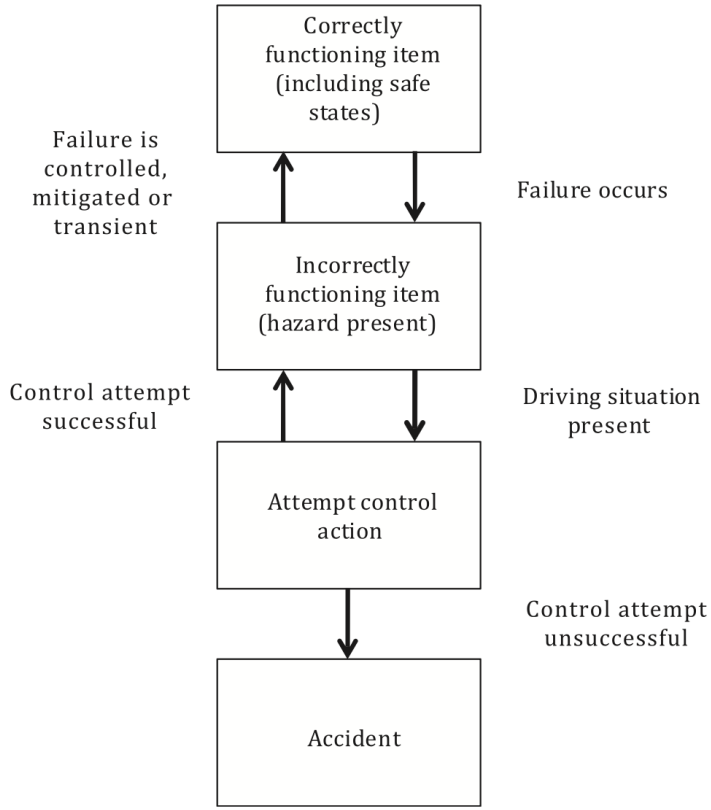


Figure 2 — State machine model of automotive risk

ISO 26262

Table 1 — Safety goals resulting from the same hazard in different situations

Failure mode	Hazard	Specific situation	Hazardous event	Possible consequences	ASIL	Safety goal	Safe state
Unintended parking brake activation	Unexpected deceleration	High speed **OR** taking a bend **OR** low friction surface	Unexpected deceleration at high speed **OR** taking a bend **OR** low friction surface	Loss of vehicle stability	*Higher ASIL*	Avoid activating the parking function without the driver's request when the vehicle is moving	EPB disabled
Unintended parking brake activation	Unexpected deceleration	Medium-low speed **AND** high friction surface	Unexpected deceleration at medium-low speed **AND** high friction surface	Rear end collision with the following vehicle	*Lower ASIL*	Avoid activating the parking function without the driver's request when the vehicle is moving	EPB disabled

PD ISO/IEC TR 5469:2024

PD ISO/IEC TR 5469:2024

BSI Standards Publication

Artificial intelligence — Functional safety and AI systems

PD ISO/IEC TR 5469:2024

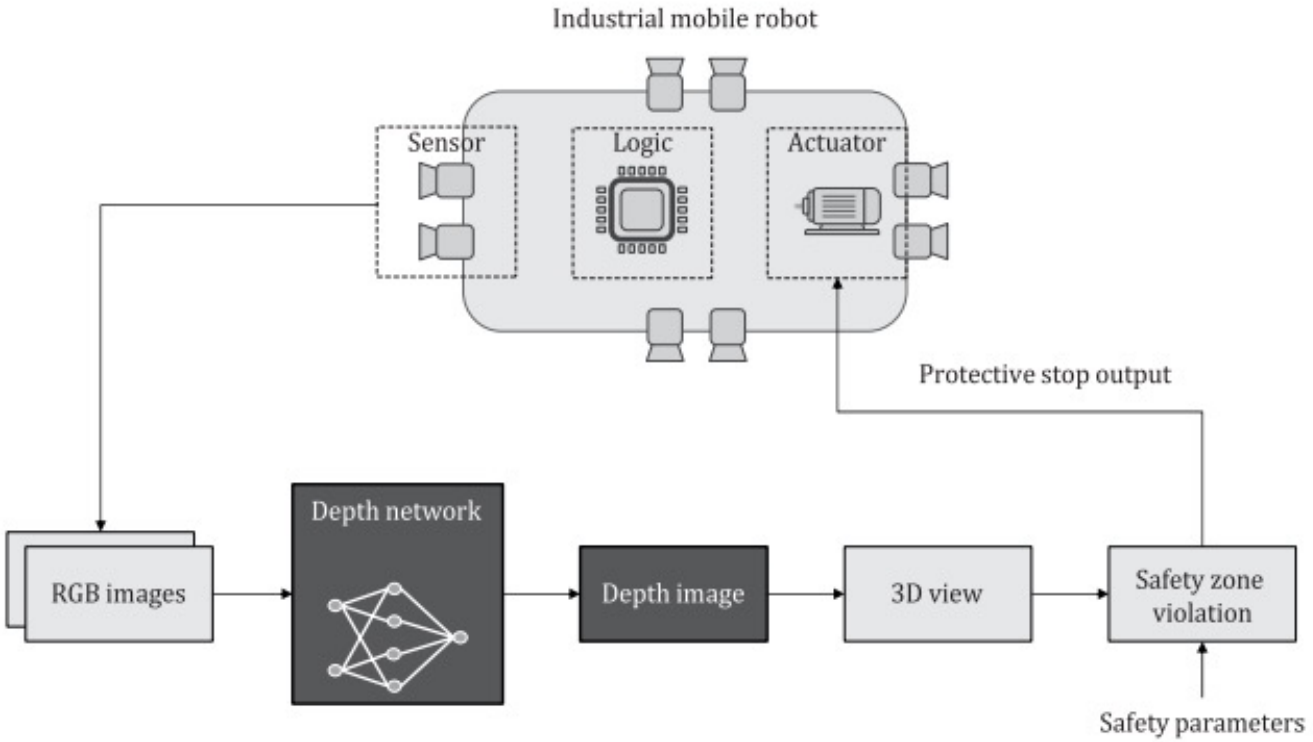


Figure B.2 — Example of industrial mobile robot

PD ISO/IEC TR 5469:2024

Based on the principles described in Clause 8, the properties addressed by the AI components are:

— Specifiability: What are the requirements of the network? How do those requirements map to existing International Standards for safety sensors, such as IEC 61496-1[144] and IEC TS 62998-1[143]? What constitutes the training images for the neural network, how are those images mapped to the operating environment? How many images, across different classes, are sufficient for training?

— Domain shift: What if the deployment environment is different from the environment used during training?

— Verifiability: How is the neural network performance assessed? How does this assessment map to existing International Standards for safety sensors, such as IEC 61496-1[144] and IEC TS 62998-1[143]?

— Robustness: How robust is the neural network to perturbation of the input data due to different causes (hardware, environmental factors, operational changes, ageing, etc.)?

— Interpretability: Are the results produced by the network understandable? Do the produced results correspond to the expected results, as defined by the safety requirements?

— Transparency: Are the components that make up the machine learning model understood? Is there a reason for design choices? Do those choices map to input requirements?

PD ISO/IEC TR 5469:2024

Table B.3 — Mapping of properties to the realization principle stages

	Acquisition from inputs or data	Knowledge induction from training data and human knowledge	Processing and generation of outputs
Specifiability	X	X	X
Domain shift	-	X	X
Verifiability	-	X	X
Robustness	X	-	X
Interpretability	-	X	X
Explainability	-	X	-

PD ISO/IEC TR 5469:2024

Table B.4 — Example property analysis

Stage: processing and generation of outputs		
Desirable property: verifiability		
Topic	**Details**	**Compliance criteria**
How is the neural network performance assessed?	— For a given input, definition of what constitutes a "correct" output by the network. — Definition of what range of inputs is evaluated.	— Pixel level KPIs. — Image level KPIs. — Sequence level KPIs. — Data set level KPIs.
How does the network performance map to existing safety International Standards and metrics?	— Mapping of measured network performance criteria to existing standard criteria. — Requirement tracing from standards performance requirements to network requirements.	— Requirement traceability or mapping documents.
How to determine verification process is accurate (e.g. unexpected behaviour due to combination of factors that cannot be seen by testing across single factors)?	— Single-dimensional vs multi-dimensional testing. — Statistical analysis of random variate testing. — Independent verification process. — Evaluation and or certification of verification tools	— Test plans. — Independently reviewed results. — Statistical analysis. — Tool qualification — Process FMEA.
How to determine when verification is complete?	— Amount of verification data — Type of verification data and how it's split into relevant parameters — Frequency with which verification is carried out	— Test plans — Predetermined stopping criteria — Process FMEA