A.1 Catalogue of safety-relevant sub-attributes

Sub-attribut	Definition (source)	safety-relevance
Functional Completeness	Capability of a product to provide a set of functions that covers all the specified tasks and intended users' objectives (ISO/IEC 25010:2023).	Missing functions may lead to unacceptable risk.
Functional Correctness	Capability of a product to provide accurate results when used by intended users (ISO/IEC 25010:2023).	Incorrect results may lead to unacceptable risk.
Functional Adaptability	Degree to which an AI system can accurately acquire information from data, or the result of previous actions, and use that information in future predictions (ISO/IEC 25059).	Continuous learning may lead to unsafe development.
Faultlessness	Capability of a product to perform specified functions without fault under normal operation (ISO/IEC 25010:2023).	A single fault may cause fatal consequences, especially in high-risk systems.
Availability	Capability of a product to be operational and accessible when required for use (ISO/IEC 25010:2023).	System breakdown may lead to physical harm.
Fault Tolerance	Capability of a product to operate as intended despite the presence of hardware or software faults (ISO/IEC 25010:2023).	Faults are inevitable and may compromise safety if not tolerated.
Recoverability	Capability of a product in the event of an interruption or a failure to recover the data directly affected and re-establish the desired state of the system (ISO/IEC 25010:2023).	Long recovery times may cause higher uncontrollable risk; environmental dependence may increase safety concerns.
Robustness	Functional correctness under any circumstances (ISO/IEC 25059).	Lack of robustness may cause failure in high-risk environments.
Self monitoring	The extent to which the system is aware of its state so it can respond appropriately to avoid going to a harmful state [1].	Inadequate reaction to hazardous situations may cause safety risks.
Time Behaviour	Capability of a product to perform its specified function under specified conditions so that the response time and throughput rates meet the requirements (ISO/IEC 25010:2023).	Delayed reaction time may cause hazards.
Resource Utilization	Capability of a product to use no more than the specified amount of resources to perform its function under specified conditions (ISO/IEC 25010:2023).	Hazard due to excessive resource consumption.
Capacity	Capability of a product to meet requirements for the maximum limits of a product parameter (ISO/IEC 25010:2023).	Hazard due to excessive resource consumption.
Co Existance	Capability of a product to perform its required functions efficiently while sharing a common environment and resources with other products, without detrimental impact on any other product (ISO/IEC 25010:2023).	Shared resource usage may impair safety-critical functions.
Interoperability	Capability of a product to exchange information with other products and mutually use the information that has been exchanged (ISO/IEC 25010:2023).	Synchronization issues may cause incorrect or delayed data transfer.

Sub-attribut	Definition (Source)	Safety-relevance
User Assistance	Capability of a product to be used by people with the widest range of characteristics and capabilities to achieve specified goals in a specified context of use (ISO/IEC 25010:2023).	Missing or insufficient assistance may lead to harm.
Operability	Capability of a product to have functions and attributes that make it easy to operate and control (ISO/IEC 25010:2023).	Complicated operability of safety functions may lead to risk.
User Error Protection	Capability of a product to prevent operation errors (ISO/IEC 25010:2023).	Lack of error protection may cause user mistakes.
Learnability	Capability of a product to have specified users learn to use specified product functions within a specified amount of time (ISO/IEC 25010:2023).	Excessive learning difficulty may lead to risks.
User Controllability	Human or another external agent can intervene in its functioning in a timely manner (ISO/IEC 25059).	Lack of intervention may lead to risk.
User Transparency	Degree to which the functionalities of the system are clear to the intended user [1].	Lack of transparency may prevent recognition of safety problems.
Documentability	See ISO/IEC/IEEE 24765.	Traceability and auditability of safety-relevant requirements.
Modularity	Capability of a product to limit changes to one component from affecting other components (ISO/IEC 25010:2023).	Lack of modularity may cause defects in one part to trigger failures in others.
Reusability	Capability of a product to be used as assets in more than one system, or in building other assets (ISO/IEC 25010:2023).	Unverified reuse may cause safety risk in other products.
Analysability	Capability of a product to be effectively and efficiently assessed regarding the impact of an intended change to one or more of its parts, to diagnose it for deficiencies or causes of failures, or to identify parts to be modified (ISO/IEC 25010:2023).	Timely and correct failure analysis ensures safety issues are identified.
Modifiability	Capability of a product to be effectively and efficiently modified without introducing defects or degrading existing product quality (ISO/IEC 25010:2023).	Changes may affect safety-relevant system parts.
Testability	Capability of a product to enable an objective and feasible test to be designed and performed to determine whether a requirement is met (ISO/IEC 25010:2023).	Ensures verification of safety-critical requirements.
Adaptability	Capability of a product to be effectively and efficiently adapted for or transferred to different hardware, software or other operational or usage environments (ISO/IEC 25010:2023).	Uncontrolled adaptation may cause safety risk.
Installability	Capability of a product to be effectively and efficiently installed successfully and/or uninstalled in a specified environment (ISO/IEC 25010:2023).	Incorrect installation or uninstallation may cause safety issues.

Sub-attribut	Definition (Source)	Safety-relevance
Replaceability	Capability of a product to replace another specified product for the same purpose in the same environment (ISO/IEC 25010:2023).	Lack of replacability may prevent safety enhancement.
Scaleability	Capability of a product to handle growing or shrinking workloads or to adapt its capacity to handle variability (ISO/IEC 25010:2023).	Scalability may affect AI system performance and safety.
Operational Contstraint	Capability of a product to constrain its operation to within safe parameters or states when encountering operational hazard (ISO/IEC 25010:2023).	Failure to maintain safe state during malfunction may cause unacceptable risk.
Risk Identification	Capability of a product to identify a course of events or operations that can expose life, property or environment to unacceptable risk (ISO/IEC 25010:2023).	Failure to identify risks may compromise safety.
Fail Safe	Capability of a product to automatically place itself in a safe operating mode, or to revert to a safe condition in the event of a failure (ISO/IEC 25010:2023).	Ensures safety in case of failure.
Hazard Warning	Capability of a product to provide warnings of unacceptable risks to operations or internal controls so that they can react in sufficient time to sustain safe operations reverting from green to yellow or red (ISO/IEC 25010:2023).	Absence of warnings may compromise safety.
Safe Intergration	Capability of a product to maintain safety during and after integration with one or more components (ISO/IEC 25010:2023).	Improper integration may compromise safety.
Data Completeness	The degree to which subject data associated with an entity has values for all expected attributes and related entity instances in a specific context of use (ISO/IEC 25012).	Incomplete data may cause unintended hazards.
Currentness	The degree to which data has attributes that are of the right age in a specific context of use (ISO/IEC 25012).	Outdated data may cause incorrect adaptation.
Consistency	The degree to which data has attributes that are free from contradiction and are coherent with other data in a specific context of use. It can be either or both among data regarding one entity and across similar data for comparable entities (ISO/IEC 25012).	Contradictory data may cause unintended hazards.
Accuracy	The degree to which data has attributes that correctly represent the true value of the intended attributes of a concept or event in a specific context of use. It has two main aspects (ISO/IEC 25012).	Syntactic or semantic inaccuracies may cause safety risks.
Data Accessibility	The degree to which data can be accessed in a specific context of use, particularly by people who need supporting technology or special configuration because of some disability (ISO/IEC 25012).	Unavailable data may cause unacceptable risk.
Credibility	The degree to which data has attributes that are regarded as true and believable by users in a specific context of use (ISO/IEC 25012).	Incorrect AI risk assessment may cause safety issues.

Sub-attribut	Definition (Source)	Safety-relevance
Compliance	The degree to which data has attributes that adhere to standards, conventions or regulations in force and similar rules relating to data quality in a specific context of use (ISO/IEC 25012).	Non-compliance may cause violation of safety norms.
Data Confidentiality	The degree to which data has attributes that ensure that it is only accessible and interpretable by authorized users in a specific context of use (ISO/IEC 25012).	Unauthorized access to safety-relevant components may cause risk.
Efficiency	The degree to which data has attributes that can be processed and provide the expected levels of performance by using the appropriate amounts and types of resources in a specific context of use (ISO/IEC 25012).	Insufficient resources may cause latency or system instability.
Precision	The degree to which data has attributes that are exact or that provide discrimination in a specific context of use (ISO/IEC 25012).	Inaccuracies may cause misinterpretation of the situation.
Data Traceability	The degree to which data has attributes that provide an audit trail of access to the data and of any changes made to the data in a specific context of use (ISO/IEC 25012).	Lack of traceability may hinder identification of systematic risks.
Understandability	The degree to which data has attributes that enable it to be read and interpreted by users, and are expressed in appropriate languages, symbols and units in a specific context of use (ISO/IEC 25012).	Lack of understandability may impair AI–human interaction in emergencies.
Availability	The degree to which data has attributes that enable it to be retrieved by authorized users and/or applications in a specific context of use (ISO/IEC 25012).	Unavailability may hinder real-time decision making.
Data Portability	The degree to which data has attributes that enable it to be installed, replaced or moved from one system to another preserving the existing quality in a specific context of use (ISO/IEC 25012).	Loss of data quality may impact safety.
Data Recoverability	The degree to which data has attributes that enable it to maintain and preserve a specified level of operations and quality, even in the event of failure, in a specific context of use (ISO/IEC 25012).	Lack of recoverability may compromise continuous operability.
Representativeness	The distribution of data corresponds to the information in the environment of the phenomenon to be captured; it is free of biases (ISO/PAS 8800).	Biased data may cause discriminatory algorithms and risks.
Independence	The datasets sufficiently avoid leakage of information amongst themselves with respect to data sources and the methods used to capture, gather, generate and process the data (ISO/PAS 8800).	Lack of data isolation (training, validation, test) may cause risk.
Data Integrity	The data are not altered by natural phenomenon (e.g. noise) or intentional action (e.g. usage of lossy data compression without consideration of impact to model, poisoning) (ISO/PAS 8800).	Noise may cause safety risks.

Sub-attribut	Definition (Source)	Safety-relevance
Temporality	The data gives sufficient consideration to time-based characteristics (e.g. timeliness, ageing, lifetime, time contributing to distribution shift) (ISO/PAS 8800).	Outdated or unsynchronized data may cause safety risk.
Intervenability	Degree to which an operator can intervene in an AI system's functioning in a timely manner to prevent harm or hazard (ISO IEC 25059).	Timely intervention may prevent harm by restoring a safe state.
Accountability	Capability of a product to enable actions of a human to be traced uniquely to the human [1].	Enables chain of responsibility.
Monitorability	The extent to which relevant indicators of an AI system are effectively observed/monitored and integrated in the operation of the system [1].	Lack of monitoring may delay failure detection and risk mitigation.
Interpretability	The extent to which the inner workings of the AI system can be analyzed in order to understand why it behaves the way it does [1].	Lack of interpretability may impair comprehension of safety decisions.
Traceability	The extent to which there exists data and processes that can record the system's decisions and link artifacts at different stages [2].	Traceability enables analysis in safety situations.
Explainability	Property of an AI system to express important factors influencing the AI system results in a way that humans can understand (ISO/IEC 22989).	Lack of explainability may hinder validation of safety decisions.

Literatur

[1] J. Kelly, S. A. Zafar, L. Heidemann, J.-V. Zacchi, D. Espinoza, and N. Mata, "Navigating the eu ai act: A methodological approach to compliance for safety-critical products," Mar. 2024.

[2] G. Li, B. Liu, and H. Zhang, "Quality attributes of trustworthy artificial intelligence in normative documents and secondary studies: A preliminary review," *Computer*, vol. 56, no. 4, pp. 28–37, Apr. 2023.