October 15, 2024

MISSION KI

Glossary v1.0

c /o acatech
German Academy of Science and Engineering

Office: Karolinenplatz 4,
80333 Munich Germany
www.acatech.de

Gefördert durch:

Bundesministerium
für Digitales
und Verkehr

aufgrund eines Beschlusses
des Deutschen Bundestages

MISSION KI
Nationale Initiative für
Künstliche Intelligenz
und Datenökonomie

TABLE OF CONTENTS

Glossary v1.0

00 Preliminary remark on v1.0

This glossary reflects the current, provisionally final state of work in the sense of a v1.0. Accordingly, the definitions are to be understood as working definitions that may be supplemented or modified as the project progresses (as part of the use case testing).

01 Generic terms

Indicator
It is not normally possible to measure directly whether individual →criteria are fulfilled. To check this, the →criteria are broken down further into indicators. Indicators result from conditions that contribute to the fulfillment of overarching →criteria at an abstract level. They thus provide
Indicators Information on specific properties of an AI system that are decisive for the qualitative fulfillment/non-fulfillment ("anchor indicators") or the degree of fulfillment of a criterion [VDE SPEC 90012: 2.28].

Observable
In order to evaluate the fulfillment of the →indicators, observables are defined that indicate in levels the extent to which an →indicator is fulfilled. The levels of the observables are predefined here and show the
Target achievement level. An observable is therefore a measurable quantity that is used to determine the status or properties of a system using an indicator [VDE SPEC 90012: 2.36].
Examples: Quality of data record documentation or risk avoidance.

Quality dimension
A quality dimension denotes a desirable, high-level, general property of an →AI system, the existence of which can be tested indirectly - i.e. via concretizations and operationalizations (→criterion, →indicator, →observable) - and which, in combination with other properties of equal rank, defines the quality of this →AI system as a whole if there is a corresponding need for protection. The model consisting of quality dimension, →criterion, →indicator, →observable is used for
The term "value" as used in the literature is referred to as "VCIO" ("Value Criterion Indicator Observable"). The term "value" used there is replaced by "quality dimension" within MISSION KI with largely the same meaning.

Criterion
Criteria are specified in order to determine whether a →KI system fulfills individual →quality dimensions. Criteria thus represent a concretization of a quality dimension in the direction of operationalization.
and observable specific facts as well as concrete risks/needs for protection.

Test
An attempt to determine certain characteristics, performance or comparable characteristics.

02 Individual quality dimensions

(AI-specific) cyber security
Resilience of →AI models, →AI components and →AI systems against AI-specific malicious external intrusions and manipulations that take place via general telecommunication networks.

Data management
Handling, processing and securing the data used within the life cycle of a →KI system.

Data quality
Property of the training, validation and test data of an →KI system with regard to its factual correctness, completeness and freedom from unjustified bias.

Data protection
Preservation and inaccessibility of certain personal data documented as information.

Explainability
Characteristics of an →KI system with regard to the basic comprehensibility and comprehensibility of functionality, behavior and output for human specialists, but also for affected persons
and users. Explainability is often understood as a property of an →AI model or →AI system that is measured locally and independently of the system design ("post-hoc").

Interpretability
Property of an → AI model that its model parameters, weights or other (mathematical) properties are as directly comprehensible as possible and directly understandable for specialist personnel. Interpretability is often explicitly given as part of the model architecture design, in contrast to the explicit choice of opaque "black box" models (e.g. the choice of an interpretable decision tree instead of a neural network for a classification task).

Performance
Property of an →KI system, in terms of its ability to achieve its specified goals and purposes as fully as possible.

Human Oversight and Control
Property of an →AI system, including its embedding in the application context, with regard to the possibility for a - technically competent - human individual to adequately observe and change the behavior and/or functioning of this →AI system in principle and during operation and, if necessary, to terminate it.

Monitoring
Procedure in which deviations between observable actual states and the desired target states are detected during the operation of an →KI system.

Non-discrimination
Characteristic of an open process carried out by an →KI system if, in the course of this process, several human individuals are treated in comparison with each other and this process is carried out in an open process.
is legally free from the mistreatment of a human individual on the basis of a legally protected characteristic.

User information
Characteristics of an →AI system with regard to the quality of information, interaction and operation by a user, including knowledge of the involvement of AI, barriers, and the quality of the user experience.
freedom and with a view to preventing nudging.

Robustness
Ability of an →AI system to maintain its regular and usual behavior and functioning in the best possible way even in the event of non-malicious, adverse, disruptive or faulty inputs or external influences.
to keep.

Traceability
Property of an →KI system with regard to the ability to record the consecutive sequence of all decisions that enter or have entered an →KI system along the entire life cycle.

Transparency
Property of an →KI system that is explainable and comprehensible. In the context of this quality standard, "transparency" also includes documentation of the properties of the →KI system.

Reliability
Property of an →KI system that has sufficient →performance, sufficient →robustness and allows sufficient →monitoring.

03 Horizontal concepts

Documentation
Systematic recording, collection, storage and provision of various types of information that comply with legal, internal or external requirements.

Fairness
Characteristic of an open process carried out by an →KI system if, in the course of this process, several human individuals are treated in comparison with each other and this process is carried out in an open process.
is free, from a legal point of view, from the mistreatment of a human individual on the basis of a legally protected characteristic and also corresponds to the ideas of justice of the individuals to be named.

Sustainability
constitution of an →AI system over its entire life cycle, with regard to the situation that the influence of the →AI system on its natural environment, including the human living environment, is not a factor. The life cycle of the living space is designed in such a way that - *ceteris paribus* - this life cycle is also possible at future points in time under at least equivalent initial and general conditions.

Safety
characteristic of a →KI system, with regard to the safety of the system for human individuals in terms of the risks to life, limb and health and for property in terms of damage in the
intended, functional regular operation.

Protection of fundamental rights
Ensuring the non-violation of alienable rights of all human individuals in the context of the use of →AI systems.

Security
Resistance of a →KI system to malicious external interventions and manipulations.

04 Further terms

Connectivity
Characteristic of the →quality standard depending on compatibility and consistency with other AI regulations such as the European AI Regulation.

Area of application
The entirety of possible input data that is relevant for an →KI system. It includes the contexts in which an →KI system can be applied or used. In certain contexts, the application domain can be used synonymously with the term "Operational Design Domain" (ODD) to define the specific conditions and parameters under which an →AI system works effectively.

Automation
The sequence of a process with little or no human intervention and with the aid of technical testing tools.

Load capacity
Property of the quality standard depending on the depth of validation and objectivity of the testing, whereby the degree of confirmation, external audits and comparable factors increase this accordingly.

Affected persons
Natural persons who are influenced or affected by an →AI system without necessarily interacting directly or actively with the system. The difference to →users lies in the type of interaction: while →users operate the →AI system directly and actively, for example by
As end users or users, affected persons have a more indirect and passive relationship with the →AI system, but can be influenced by its decisions or functions.

Efficiency
Characteristic of the quality standard depending on time, personnel and financial expenditure while at the same time ensuring a high level of quality.

Protected properties
Protected characteristics are (based on the AGG): Race, ethnic origin, gender, religion or belief, disability, age, sexual identity.

AI provider (AI provider)
A "provider" is a natural or legal person, public authority, institution or other body that

develops or has developed a →KI system or a →KI model with a general purpose and places it on the market under its own name or trademark or puts the →KI system into operation under its own name or trademark, whether in return for payment or free of charge
ing. [cf. Art. 3 para. 3 AI Regulation].

AI Deployer

An "operator" is a natural or legal person, public authority, agency or other body which uses an →KI system under its own responsibility, unless the →KI system is used in the course of a personal and non-professional activity. [see Art. 3 para. 4 AI Regulation].

AI component

An "AI component" comprises an implemented →AI model; possibly together with the methods that relate directly to the pre- or post-processing of the inputs/outputs of this model and
their interfaces.
Example: The aforementioned AI image recognition model including methods for pre-processing the images, which takes raw image/video data as input and makes a statement as output as to whether a person can be seen in the image.

AI life cycle (AI life cycle)

Development of a system, product, service, project or other human-made entity - that uses AI - from conception to decommissioning. [based on and continuation of ISO/IEC 22989:2022]

AI model

An "AI model" only includes the functional, AI-specific parameters, the weights and biases (inferential input-output mappings) and the architecture, if applicable; this does not include the further development of the AI model.

leading implementation and integration, this is only covered by the term → AI component
will.
Example: A neural network for image processing that receives numerical values assigned to pixels as input and outputs the probability that a person can be seen in the image.

AI system

An "AI system" is a machine-based system that is designed to operate with varying degrees of autonomy and that, once operational, can be adaptive and that derives from inputs received for explicit or implicit goals how to produce outputs such as predictions, insights, recommendations or decisions that affect physical or virtual environments.
can influence. [For the purposes of this quality standard, an "AI system" is defined in technical terms as a functional combination of one or more → AI component(s) and non-AI components with regard to a specific purpose and a specific purpose.
a concrete application context.

AI developer
Combination of →AI provider and →AI operator; providers who further develop an AI system on the basis of a given AI system (e.g. an LLM).

Network
Connection of at least two computers or other electronic devices that enables the exchange of data and the use of shared resources.

Low threshold
Offer or service characterized by a low required effort (e.g. short test duration) with a high benefit at the same time (high quality ensured).

Users
Natural persons who interact directly and actively with an →KI system, either as end users who use the →KI system for personal or business purposes, or as users who use the →KI system in a professional context. The term includes both those who use the
→KI system, as well as those who use it to achieve specific goals. Depending on the context, this may involve the use of →AI systems in everyday applications or in specialized, professional applications. scenarios.

Proprietary data
Data that is assigned under proprietary law, including licensed and copyrighted data.

Test method
Methodical procedure for collecting individual or several related inspection events and their evaluation in the context of the inspection procedure. An inspection method can be both manual and
also contain automated parts (→test workflow).

Test platform
Technical framework (software) that is used in perspective to execute →test workflows. The purpose of the test platform is to support and implement the technical components of the test system.
(connection to the project: test center/examiners use test platform as part of the test set), which provide adequate testing tools and other aids and resources for conducting examinations.

Test depth
Characteristic of the audit, which is composed of the degree of confirmation regarding the truthfulness and plausibility of the audit documents and the degree of detail according to which evidence is obtained.

Test environment
A test environment provides the (physical) infrastructure and resources for testing one or more specific special test objects provided and has for this access to one instance of the
→test platform. It is part of a →test center.

Test procedure
Procedure within the scope of the audit, which includes determining the depth of the audit, the methodology for obtaining evidence, the comparability and the evaluation of the criteria.

Test tool
Algorithms (software) that can be used (by developers, testers) to create a new algorithm for a given →KI system or several →KI components to obtain measurable and, if possible, verifiable evidence with regard to a specific test criterion. Tools are used to facilitate a
Audit like an automated questionnaire is not considered an audit tool.

Quality standard
A standard is a document that defines a procedure so that it is objectively verifiable whether a test object conforms to the criteria defined in the standard. A quality standard is a standard whose criteria allow conclusions to be drawn about the quality of the test object.

Rating criteria
Part of the assessment methodology that aggregates the individual →criteria into an overall assessment.

Test environment
Technical and organizational infrastructure required to carry out a technical test. It includes hardware, instrumentation, simulators, software tools and other supporting resources.

Test workflow/pipeline
The algorithmic flow chart describes an automated component for collecting individual or multiple test evidence as part of a → test method using one or more → test tools, including their respective configuration.

Test center
Institutional or organizational unit that carries out tests in one or more →test environments (test benches). The →test environments are sufficiently adapted to specific use cases and the applications or systems to be tested.

Comparability
Characteristic of the →quality standard depending on the uniformity of the tests and the replicability of the tests with the aim of objectification.

Accessibility
Property of either a test statement, which provides information about its comprehensibility, or of a test approach, which describes the extent of applicability to specific target groups.