ITUPublications

International Telecommunication Union
Standardization Sector

ITU-T Focus Group Deliverable

(09/2022)

Focus Group on Artificial Intelligence for Health
(FG-AI4H)

FG-AI4H DEL0.1

Common unified terms in artificial intelligence for health

ITU-T FG-AI4H Deliverable

DEL0.1 – Common unified terms in artificial intelligence for health

Summary

This deliverable contains a glossary with agreed terminology in artificial intelligence (AI) for health for use not only across the various FG-AI4H Deliverables, but also to promote the harmonized use of important AI for health terms across the different disciplines involved in this cross-disciplinary field.

Keywords

Artificial intelligence, glossary, health, medical devices, terminology.

Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Change Log

This document contains Version 1 of the Deliverable DEL0.1 on "*Common unified terms in artificial intelligence for health*", approved by the FG-AI4H after a two week consultation following its meeting P in Helsinki, 19-22 September 2022.

Editor: Markus Wenzel E-mail: markus.wenzel@hhi.fraunhofer.de
Fraunhofer HHI
Germany

Contributors: (in alphabetical order, by family name)

Shada Alsalamah E-mail: alsalamahs@who.int
World Health Organization
Switzerland

Xiaomi An E-mail: anxiaomi@ruc.edu.cn
Renmin University of China
China

Pat Baird E-mail: pat.baird@philips.com
Philips
USA

Simão Campos E-mail: simao.campos@itu.int
International Telecommunication Union
Switzerland

Henry Hoffmann
Ada Health GmbH
Germany
E-mail: henry.hoffmann@ada.com

Joachim Krois
dentalXrai
Germany
E-mail: joachim.krois@dentalxr.ai

Stephanie Kuku
World Health Organization
Switzerland
E-mail: kukus@who.int

Rohit Malpani
World Health Organization
Switzerland
E-mail: malpanir@who.int

Sameer Pujari
World Health Organization
Switzerland
E-mail: pujiaris@who.int

Andreas Reis
World Health Organization
Switzerland
E-mail: reisa@who.int

Eva Weicken
Fraunhofer HHI
Germany
E-mail : eva.weicken@hhi.fraunhofer.de

Table of Contents

Technical Report ITU-T FG-AI4H Deliverable

DEL0.1 – Common unified terms in artificial intelligence for health

1 Introduction

The International Telecommunication Union (ITU) is the UN specialized agency for information and communication technologies. The World Health Organization (WHO) is the UN specialized agency for health. Both organizations collaborated to establish an open group of experts to develop a generalizable benchmarking framework for health solutions based on artificial intelligence (AI), the ITU/WHO Focus Group on AI for Health (FG-AI4H).

A glossary with agreed terminology for the field of artificial intelligence for health can be of great help for many actors in this interdisciplinary area. Common unified terms and definitions may foster the dialogue between experts of different professional backgrounds such as clinicians, developers, machine learning scientists, regulators, ethicists and public health officials. This document contains an initial collection of selected terms and definitions to be extended later as needed. The document adopts terms and definitions from both the scientific literature and authoritative sources when available, and provides new definitions in the health AI context. The document also provides a collection of acronyms and abbreviations typically in the field, and a bibliography with all references.

The design principles for the document are as follows. The authors differentiate between "terms defined elsewhere" in the sections x.1 and "terms defined here" in the sections x.2. Where a definition originates from elsewhere, e.g., from standards or scientific literature, this definition is taken as given and quoted verbatim, as far as possible, remaining faithful to the original. In the case of concerns regarding a definition, either a better definition from a different literature source can be identified or the definition amended under "terms defined here" in section x.2. In order to promote synergy with other groups, the authors first try to rely on existing definitions from reputable literature/standardization sources.

2 Technical terms and definitions

2.1 Terms defined elsewhere

This glossary adopts the following technical terms defined elsewhere:

2.1.1 algorithm [IMDRF/SaMD-N41]: A finite set of instructions (or rules) that defines a sequence of operations for solving a particular computational problem for all problem instances for a problem set.

2.1.2 application [ITU-T H.764]: A functional implementation realized as software running in one or spread over several interplaying hardware entities.

2.1.3 artificial intelligence (AI):

2.1.3.1 [Nilsson 1998]: AI, broadly (and somewhat circularly) defined, is concerned with intelligent behaviour in artefacts. Intelligent behaviour, in turn, involves perception, reasoning, learning, communicating, and acting in complex environments.

NOTE – The definition of [Nilsson 1998] shall serve as a starting point in this draft glossary for exploring, collecting, reconciling, and consolidating different definitions of the term "AI", which is subject of controversial discussions.

2.1.3.2 [IMDRF/AIMD-N67]: AI is a branch of computer science, statistics, and engineering that uses algorithms or models to perform tasks and exhibit behaviors such as learning, making decisions and making predictions. The subset of AI known as Machine Learning (ML) allows ML models to

be developed by ML training algorithms through analysis of data, without models being explicitly programmed.

2.1.3.3 [ISO/IEC 22989]: Capability of an engineered system to acquire, process and apply knowledge.

2.1.3.4 [ISO/IEC 22989]: <discipline> research and development of mechanisms and applications of AI systems.

NOTE 1 – Research and development can take place across any number of fields such as computer science, data science, humanities, mathematics and natural sciences.

2.1.4 artificial intelligence system (AI system):

2.1.4.1 [ISO/IEC 22989] engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives.

NOTE 1 – The engineered system can use various techniques and approaches related to artificial intelligence to develop a model to represent data, knowledge, processes, etc. which can be used to conduct tasks.

NOTE 2 – AI systems are designed to operate with varying levels of automation.

2.1.4.2 [EC 2021]: Means software that is developed with one or more of the techniques and approaches[1] and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with.

2.1.5 autonomy/autonomous [ISO/IEC 22989]: Characteristic of a system that is capable of modifying its intended domain of use or goal without external intervention, control or oversight.

2.1.6 batch learning [CTA-2089]: Method of training where examples are presented in groups to the model. Typically used when a large amount of pre-recorded data.

2.1.7 bias [ISO/IEC 22989]: Systematic difference in treatment of certain objects, people, or groups in comparison to others.

NOTE – Bias is used both in a technical/statistical context and in ethical/legal discussions with different definitions. See clause 5.1.6 for the ethics perspective on this term.

2.1.8 class-activation map [SPIRIT-AI]: Class-activation maps are particularly relevant to image classification AI interventions. Class-activation maps are visualisations of the pixels that had the greatest influence on predicted class, by displaying the gradient of the predicted outcome from the model with respect to the input. They are also referred to as "saliency maps" or "heat maps".

2.1.9 continuous learning [ISO/IEC 22989]: Incremental training of an AI system that takes place on an ongoing basis during the operation phase of the AI system life cycle.

2.1.10 convolutional neural networks (CNN) [Deep Learning]: Convolutional networks, also known as convolutional neural networks, are a specialized kind of neural network for processing data that has a known grid-like topology.

2.1.11 deep learning (DL) [ISO/IEC 22989]: Approach to creating rich hierarchical representations through the training of neural networks with many hidden layers.

NOTE – Deep learning is also known as deep neural network learning.

2.1.12 fine-tuning [SPIRIT-AI]: Modifications or additional training performed on the AI intervention model, done with the intention of improving its performance.

---

[1] AI techniques and approaches: (a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; (b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; (c) Statistical approaches, Bayesian estimation, search and optimization methods.

2.1.13 genetic algorithm (GA) [ISO/IEC 22989, 3.1.14]: Algorithm which simulates natural selection by creating and evolving a population of individuals (solutions) for optimization problems.

2.1.14 heteronomy/heteronomous [ISO/IEC 22989]: Characteristic of a system operating under the constraint of external intervention, control or oversight.

2.1.15 locked algorithm [FDA 2019]: An algorithm that provides the same result each time the same input is applied to it and does not change with use.

2.1.16 machine learning (ML):

2.1.16.1[CONSORT-AI]: A field of computer science concerned with the development of models/algorithms that can solve specific tasks by learning patterns from data, rather than by following explicit rules. It is seen as an approach within the field of AI.

2.1.16.2[ITU-T Y.3172]: Processes that enable computational systems to understand data and gain knowledge from it without necessarily being explicitly programmed.

2.1.17 neural networks (NN) [CONSORT-AI]: Simplified from Artificial Neural Networks (ANN). An ANN is based on a collection of connected units or nodes called artificial neurons which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons.

2.1.18 reliability [ISO/IEC 22989]: Property of consistent intended behaviour and results.

2.1.19 Semi-supervised Machine Learning [ISO/IEC 22989]: Machine learning that makes use of both labelled and unlabelled data during training.

2.1.20 Supervised machine learning [ISO/IEC 22989]: Machine learning that makes use of labelled data during training.

2.1.21 training [ISO/IEC 22989]: Process to establish or to improve the parameters of a machine learning model, based on a machine learning algorithm, by using training data.

2.1.22 Training dataset [ISO/IEC TR 24028]: A dataset used in the training process to establish or improve the parameters of a machine learning model based on a machine learning algorithm.

2.1.23 unsupervised machine learning [ISO/IEC 22989]: Machine learning that makes use of unlabelled data during training.

2.2 Terms defined here

This glossary defines the following technical terms:

2.2.1 test dataset: A subset of the data that is never shown to the model during training, and which is used to verify that the model has learned what it was supposed to.

NOTE – Adapted from [ISO/IEC 22989], which defines "test data: data used to assess the performance of a final machine learning model. [...] Test data is disjoint from training data and validation data. [...]. The test set is used to verify that the model has learned what it was supposed to.")

2.2.2 AI technology: The term AI technology is referred to any technology (e.g., machine learning, deep learning, natural language processing, computer vision, etc) used to develop an AI system.

3 Statistical terms and definitions

3.1 Terms defined elsewhere

This glossary adopts the following statistical terms defined elsewhere:

3.1.1 area under the receiver operating characteristic curve (AUROC) [Ekelund 2012]: The area under a receiver operating characteristic curve is a measure of the usefulness of a test in general,

where a greater area means a more useful test, the areas under ROC curves are used to compare the usefulness of tests.

3.1.2 sensitivity [Yerushalmy1947]: The probability of correct diagnosis of "positive" cases.

3.1.3 specificity [Yerushalmy1947]: The probability of correct diagnosis of "negative" cases.

NOTE – This original paper by Jacob Yerushalmy from 1947 introduced the terms of sensitivity and specificity on the example of tuberculosis diagnosis from X-ray and further explains: "By 'positive' cases are meant persons who have X-ray evidence suggestive of tuberculosis, and 'negative' cases refer to persons who have no such evidence."

3.2 Terms defined here

This glossary does not provide own definitions of statistical terms.

4 Clinical & scientific & evaluation terms and definitions

4.1 Terms defined elsewhere

This glossary adopts the following evaluation terms defined elsewhere:

4.1.1 clinical data [IMDRF/MDCE-N57]: Safety, clinical performance, and/or effectiveness information that is generated from the clinical use of a medical device.

4.1.2 clinical evaluation [IMDRF/MDCE-N57]: A set of ongoing activities that use scientifically sound methods for the assessment and analysis of clinical data to verify the safety, clinical performance and/or effectiveness of the device when used as intended by the manufacturer.

4.1.3 clinical evidence [IMDRF/MDCE-N57]: The clinical data and its evaluation pertaining to a medical device.

4.1.4 clinical investigation [IMDRF/MDCE-N57]: Any systematic investigation or study in or on one or more human subjects, undertaken to assess the safety, clinical performance and/or effectiveness of a medical device.

4.1.5 clinical outcome [SPIRIT-AI]: Measured variables in a clinical trial that are used to assess the effects of an intervention.

4.1.6 clinical outcome assessment [FDA-NIH BEST]: Assessment of a clinical outcome ["clinical outcome" is in the source [FDA-NIH BEST] specifically defined as: "An outcome that describes or reflects how an individual feels, functions or survives"] can be made through report by a clinician, a patient, a non-clinician observer or through a performance-based assessment. There are four types of COAs. clinician-reported outcome, observer-reported outcome, patient-reported outcome, performance outcome.

4.1.7 clinical performance [IMDRF/MDCE-N57]: The ability of a medical device to achieve its intended clinical purpose as claimed by the manufacturer.

4.1.8 clinical trials [IMDRF/MDCE-N57]: A properly conducted clinical investigation, including compliance to the clinical investigation plan and local laws and regulations, ensures the protection of human subjects, the integrity of the data and that the data obtained is acceptable for the purpose of demonstrating the SaMD's conformity to the Essential Principles.

4.1.9 clinical validation [IMDRF/SaMD-N41]: The ability of a SaMD to yield a clinically meaningful output associated to the target use of SaMD output in with the target health care situation or condition identified in the SaMD definition statement.

4.1.10 development environment [SPIRIT-AI]: The clinical, and operational settings from which the data used for training the model are generated. This includes all aspects of the physical setting (such as geographical location, physical environment), operational setting (such as integration with

an electronic record system, installation on a physical device) and clinical setting (such as primary, secondary and/or tertiary care, patient disease spectrum).

4.1.11 effectiveness [IMDRF/MDCE-N57]: The ability of a medical device to achieve clinically meaningful outcome(s) in its intended use as claimed by the manufacturer.

4.1.12 input data [SPIRIT-AI]: The data that needs to be presented to the AI system to allow it to serve its purpose.

4.1.13 Intended use [SPIRIT-AI]: The objective intent of the manufacturer regarding the use of a product, process or service as reflected in the specifications, instructions and information provided by the manufacturer.

4.1.14 output data [SPIRIT-AI]: The predicted output given by the AI system based on processing of the input data. The output data can be presented in different forms, including a classification (including diagnosis, disease severity or stage, or recommendation such as referability), a probability, a class-activation map, etc.

4.1.15 performance error [SPIRIT-AI]: Instances in which the AI system fails to perform as expected. This term can describe different types of failures, and it is up to the investigator to specify what should be considered a performance error, preferably based on prior evidence. This can range from small decreases in accuracy (compared to expected accuracy) to erroneous predictions or the inability to produce an output, in certain cases.

4.1.16 post-market clinical follow-up study (PMCF-study) [ISO/TR 20416]: Study carried out following marketing approval intended to answer specific questions relating to clinical safety or performance (i.e., residual risks) of a medical device when used in accordance with its approved labelling.

4.1.17 post-market surveillance (PMS) [ISO 13485]: Systematic process to collect and analyse the performance of medical devices that have been placed on the market.

4.1.18 reproducibility [ISO/IEC 27037]: Property of a process to get the same test results on a different testing environment (different computer, hard drive, operator, etc.).

4.1.19 safety [IMDRF/MDCE-N57]: Acceptability of risks as weighed against benefits, when using the medical device according to the manufacturer's labelling.

4.1.20 scientific validity (valid clinical association) [IMDRF/SaMD-N41]: The extent to which the SaMD's output (concept, conclusion, measurements) is clinically accepted or well founded (based on an established scientific framework or body of evidence) and corresponds accurately in the real world to the healthcare situation and condition identified in the SaMD definition statement (corresponds to the level of clinical acceptance of the SaMD's output).

4.1.21 usability [ISO/IEC 25010]: Degree to which a product or system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

NOTE 1 – Adapted from ISO 9241-210 and ITU-T F.901 (03/1993).

NOTE 2 – Usability can either be specified or measured as a product quality characteristic in terms of its subcharacteristics, or specified or measured directly by measures that are a subset of quality in use.

4.1.22 verification [ISO 9000]: Confirmation, through the provision of objective evidence, that specified requirements have been fulfilled.

NOTE 1 – The objective evidence needed for a verification can be the result of an inspection or of other forms of determination such as performing alternative calculations or reviewing documents.

NOTE 2 – The activities carried out for verification are sometimes called a qualification process.

NOTE 3 – The word "verified" is used to designate the corresponding status.

4.1.23 validation [ISO 9000]: confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled.

NOTE 1 – The objective evidence needed for a validation is the result of a test or other form of determination such as performing alternative calculations or reviewing documents.

NOTE 2 – The word "validated" is used to designate the corresponding status.

NOTE 3 – The use conditions for validation can be real or simulated.

4.2 Terms defined here

This glossary defines the following clinical, scientific, and evaluation terms:

4.2.1 analytical validation: Evaluation of the adequacy of the artificial intelligence (AI) model

'in silico' before being implemented 'in vivo' in the clinical pathway.

4.2.2 clinical validation: Evaluation of the artificial intelligence (AI) system through interventional or clinical studies in which the whole AI health technology is evaluated in the context of the clinical pathway.

4.2.3 external validation: The process of evaluating the performance of the artificial intelligence (AI) model using previously unseen, and independent data 'in silico'.

5 Ethics terms and definitions

5.1 Terms defined elsewhere

This glossary adopts the following ethics terms defined elsewhere:

5.1.1 accountability [ISO/IEC TR 24028]: Property that ensures the actions of an entity may be traced uniquely to that entity.

5.1.2 anonymization [WHO AI-EG]: With respect to personal data, a sub-category of de-identification whereby both direct and indirect personal identifiers are removed, and technical safeguards are used to ensure zero risk of re-identification.

5.1.3 automation bias [WHO AI-EG]: A lack of consideration by a healthcare provider of whether an automated technology meets their needs or those of the patient. This may lead a provider to overlook errors that should have been spotted by human-guided decision-making.

5.1.4 autonomy [WHO GHE]: Most often taken to refer to the ability of an individual to be his or her own person, to make his/her own choices on the basis of his/her motivations, without manipulation by external forces. However, others in a more Kantian tradition see autonomy as being firmly related to accepting and acting on the basis of one's obligations, i.e., acting morally, the precise oppose of what one wants.

5.1.5 beneficence [WHO GHE]: Principle requiring that governments, health care providers, and researchers do good for, provide benefit to, or make a positive contribution to the welfare of populations, patients and study participants.

5.1.6 bias [WHO AI-EG]: A threat to inclusiveness and equity that represents a departure, often arbitrary, from equal treatment.

NOTE – Bias is used both in ethics/legal discussions and in a technical/statistical context with different definitions. See clause 2.1.7 for the technical perspective on this term.

5.1.7 biosurveillance [WHO AI-EG]: A form of surveillance for health data and other biometrics, such as facial features, fingerprints, temperature, and pulse.

5.1.8 black-box algorithms [WHO AI-EG]: Algorithms that make inferences and decisions that are not understood, including by their developers.

5.1.9 closed data [ITU-T FG-DPM D0.1]: Data that requires access control to be divulgated.

5.1.10 confidentiality [WHO GHE]: The obligation to keep information secret unless its disclosure has been appropriately authorized by the person concerned or, in extraordinary circumstances, by the appropriate authorities.

5.1.11 control problem [WHO AI-EG]: Wherein developers and designers of AI may not be held responsible, as AI guided systems function independently of their developers and may evolve in ways that the developer could claim were not foreseeable.

5.1.12 co-regulation [WHO AI-EG]: Wherein governments and companies rely on each other to assess and regulate a technology. While such models of oversight may assist governments in understanding a technology, they may limit a government's exercise of independent judgment and encourage governments to trust that companies are willing to strictly self-regulate their practices.

5.1.13 data altruism [WHO AI-EG]: Also known as data solidarity, this allows companies to collect personal and non-personal data on individuals for projects that are in the public interest.

5.1.14 data colonialism [WHO AI-EG]: Generating data from low- and middle-income countries in which the data are used for commercial or non-commercial purposes without due respect for consent, privacy, or autonomy.

5.1.15 data exchange [ITU-T FG-DPM D0.1]: Accessing, transferring and archiving of data.

5.1.16 data integrity [FDA 2021]: Data Integrity can be defined as "the completeness, consistency, and accuracy of data."

5.1.17 data portability [WHO AI-EG]: The right of individuals to obtain their personal data in a machine-readable format from one controller that can be sent to another controller.

5.1.18 data processing and management [ITU-T FG-DPM D0.1]: Combination of all activities either directly performed on or indirectly influencing data.

NOTE 1 – Directly performed activities include among others [collecting/acquiring/capturing], exchanging, storing, securing, manipulating, reusing, aggregating, curating, disposing, monetizing and deleting data.

NOTE 2 – Indirectly influencing activities include among others policy and standards making, skills and innovation enhancement.

5.1.19 data protection laws [WHO AI-EG]: Rights-based approaches that provide standards for regulating data processing that both protect the rights of individuals and establish obligations for data controllers and processors.

5.1.20 data sharing [ITU-T FG-DPM D0.1]: The process of data exchange among different parties with specified conditions.

5.1.21 data triangulation [WHO AI-EG]: Techniques that can be used to reconstruct a de-identified, incomplete dataset by a third party for re-identification of an individual.

5.1.22 de-identification [WHO AI-EG]: With respect to personal data, preventing any connection of personal identifiers to information.

5.1.23 digital divide [WHO AI-EG]: The uneven distribution of access to, use of or effect of information and communication technologies among any number of distinct groups.

5.1.24 digital welfare state [WHO AI-EG]: The use of AI to provide public services, including an assessment of whether an individual qualifies for certain services. Digital data and technologies are used to automate, predict, identify, or disqualify potential recipients of social welfare, including healthcare benefits. There is concern that the digital welfare state could undermine access to social services and welfare and especially affect poor and marginalised populations.

5.1.25 ethics [WHO GHE]: Branch of knowledge concerned with questions about right versus wrong conduct and what constitutes a good or bad life, as well as the justificatory basis for such questions.

5.1.26 explainability:

5.1.26.1[WHO AI-EG]: Improving the scientific understanding of an algorithm to understand how a system arrives at a decision. AI technologies should be explainable to the extent possible and according to the capacity of those to whom the explanation is directed. Those who might request or require an explanation should be well informed, and the educational information must be tailored to each population, including, for example, marginalized populations.

5.1.26.2[ISO/IEC 22989, 3.5.7]: Property of an AI system to express in important factors influencing the AI system results in a way that humans can understand

NOTE 1 – It is intended to answer the question "Why?" without actually attempting to argue that the course of action that was taken was necessarily optimal.

5.1.27 fairness [WHO AI-EG]: Ensuring that all persons are treated fairly, which includes the requirement to ensure that no person or group is subject to discrimination, neglect, manipulation, domination, or abuse.

5.1.28 federated data [WHO AI-EG]: A way to enable access to health data, including genomic data, that must remain inside a country or institution because of their sensitivity. Data do not leave the participating organization that holds them, but authorized users can make queries that allow them to access data, for example to train an algorithm.

5.1.29 Human rights [WHO GHE]: Fundamental freedoms and rights enshrined in a set of universal legal statements. Some of the most important characteristics of human rights are that: they are acknowledged in international declarations; states and state actors are obliged to respect them; they cannot be waived or taken away (although the enjoyment of particular human rights may be limited in exceptional circumstances); they are interdependent and inter-related; and they are universal.

5.1.30 human warranty [WHO AI-EG]: Evaluation by patients and clinicians in the development and deployment of AI technologies. Regulatory principles are applied upstream and downstream of the algorithm by establishing points of human supervision. Points of human supervision are identified by discussions among professionals, patients, and designers.

5.1.31 impact assessment [WHO AI-EG]: An impact assessment is used to predict the consequences of a current or proposed action, policy, law, regulation or, as in the case of use of AI for health, a new technology or service. Impact assessments can provide both technical information on possible consequences and risks (both positive and negative) and improve decision-making, transparency, and participation of the public and introduce a framework for appropriate follow-up and measurement.

5.1.32 inclusiveness [WHO AI-EG]: A requirement that AI is designed to encourage the widest possible appropriate, equitable use and access, irrespective of age, gender, income, ability or other characteristics.

5.1.33 informed consent [WHO GHE]: Agreement to a certain course of action, such as treatment or participation in research, on the basis of complete and relevant information by a competent individual without coercion.

5.1.34 many hands problem [WHO AI-EG]: Since the development of AI involves contributions from many agents, it is difficult, both legally and morally, to assign responsibility, which is diffused among all the contributors to the AI-guided technology.

5.1.35 nonmaleficence [WHO GHE]: A principle requiring that health care providers and researchers do not inflict undue harm, either intentionally or through negligence.

5.1.36 open data [ITU-T FG-DPM D0.1]: Any information that has been made available for anyone under a legal framework to access, alter, and share without restrictions.

NOTE – It can be from a public source, e.g., government data, or from a business, e.g., company intelligence, and can be used for both commercial and non-commercial purposes.

5.1.37 peer disagreement [WHO AI-EG]: Disagreement between two competent experts – an AI machine and a doctor, where in there is no means of combining the decisions or of reasoning with the algorithm, and no clear rules for determining who is right.

5.1.38 personal data [PAS 185]: Data which relates to a living individual who can be identified: a) from those data; or b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

5.1.39 privacy [WHO GHE]: Privacy seeks to protect a person from scrutiny by others. Respect for privacy implies that a person should not be expected to share personal information unless they so choose. Any violation of privacy requires ethical justification although it might be outweighed by other considerations in some cases (i.e., for the protection of the common good).

5.1.40 processed data [ISO 5127]: Data which have been transformed from raw data or from an earlier data stage into a more refined stage by data cleaning, sorting, linking, verifying and similar operations.

5.1.41 pseudo-anonymization [WHO AI-EG]: The processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional information provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

5.1.42 raw data [PAS 185]: Data that has not been processed for use.

5.1.43 responsibility [WHO AI-EG]: Responsibility ensures that individuals and entities are held accountable for any adverse effects of their actions and is necessary to maintain trust and to protect human rights.

5.1.44 responsiveness [WHO AI-EG]: A requirement that designers, developers and users continuously, systematically, and transparently examine an AI technology to determine whether it is responding adequately, appropriately, and according to communicated expectations and requirements in the context in which it is used.

5.1.45 sustainability [WHO AI-EG]: AI technologies that can be fully integrated and sustained in a health-care system and designed to minimize its ecological footprint and increase energy efficiency.

5.1.46 transparency:

5.1.46.1 [WHO AI-EG]: Transparency requires that sufficient information be published or documented before the design and deployment of an AI technology. Such information should facilitate meaningful public consultation and debate on how the AI technology is designed and how it should be used. Such information should continue to be published and documented regularly and in a timely manner after an AI technology is approved for use.

5.1.46.2 [ISO/IEC 22989; ISO/IEC TS 6254]: <organization> property of an organization that appropriate activities and decisions are communicated to relevant *stakeholders* (3.5.13) in a comprehensive, accessible and understandable manner.

NOTE 1 – Inappropriate communication of activities and decisions can violate security, privacy or confidentiality requirements.

5.1.46.3 [ISO/IEC 22989; ISO/IEC TS 6254]: <system> property of a system that appropriate information about the system is made available to relevant stakeholder.

NOTE 1 – Appropriate information for system transparency can include aspects such as features, performance, limitations, components, procedures, measures, design goals, design choices and assumptions, data sources and labelling protocols.

NOTE 2 – Inappropriate disclosure of some aspects of a system can violate security, privacy or confidentiality requirements.

5.1.46.4 [ISO/IEC DIS 25059]: <quality model> degree to which appropriate information about the AI system is communicated to relevant stakeholders

NOTE – Appropriate information for AI system transparency can include aspects such as features, components, procedures, measures, design goals, design choices and assumptions.

5.1.46.5 [ISO/IEC TS 5723]: <information> open, comprehensive, accessible, clear and understandable presentation of information.

5.1.46.6 [ISO/IEC TS: 5723]: <systems> property of a system or process to imply openness and accountability property of a system or process to imply openness and accountability.

5.1.47 trust [ITU-T X.1252]: The reliability and truth of information or the ability and disposition of an entity to act appropriately, within a specified context.

5.1.48 visibility [ISO/IEC 27036-1]: Property of a system or process that enables system elements and processes to be documented and available for monitoring and inspection

5.2 Terms defined here

This glossary defines the following ethics term:

5.2.1 trustworthy AI: In the context of this document it refers to artificial intelligence (AI) systems and technologies that meet stakeholder's expectation in terms of bias, explainability, provenance and other desirable characteristics. Therefore, stakeholders involved in the development, deployment, or operation of such AI-based systems should be held accountable for their proper functioning.

6 Product terms and definitions

6.1 Terms defined elsewhere

This glossary adopts the following product term defined elsewhere:

6.1.1 Software as a Medical Device (SaMD) [IMDRF SaMD-N12]: Software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device.

6.2 Terms defined here

This glossary defines the following product term:

6.2.1 total product life cycle (TPLC): A conceptual framework for holistically managing any product or service throughout all of its stages, e.g., from inception to introduction, growth, maturity, and decline.

7 Policy terms and definitions

7.1 Terms defined elsewhere

This glossary adopts the following policy terms defined elsewhere:

7.1.1 data governance [ITU-T FG DPM D0.1]: Set of activities aimed to design, implement and monitor a strategic plan for data asset management.

7.1.2 high income countries (HIC): List of countries with higher levels of income that is defined by the World Bank and reviewed regularly, as found at https://datahelpdesk.worldbank.org/knowledgebase/articles/906519.

7.1.3 low and lower-middle income countries (LMIC): List of countries with lower levels of income that is defined by the World Bank and reviewed regularly, as found at https://datahelpdesk.worldbank.org/knowledgebase/articles/906519.

8.2 Terms defined here

This glossary defines the following policy terms:

7.2.1 focus group: A group created under the provisions of [ITU-T A.7] to help advance the work of the ITU Telecommunication Standardization Sector (ITU-T) study groups and to encourage the participation of members of other standards organizations, including experts and individuals who may not be members of ITU. Focus group activities may include an analysis of gaps between current Recommendations and expected Recommendations, and provide material for consideration in the development of Recommendations. They augment an ITU-T study group work programme by providing an alternative working environment for the quick development of specifications in their chosen areas. (Adapted from [ITU-T A.1].)

7.2.2 sustainable development goals (SDGs) (Adapted from [UN-SDGs]): The Sustainable Development Goals of the United Nations are urgent "calls for action" by all countries in 17 areas of human development in a global partnership. They recognize that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality, and spur economic growth – all while tackling climate change and working to preserve our oceans and forests. SDG 3 "Ensure healthy lives and promote well-being for all at all ages" is particularly relevant for AI for Health.

8 Abbreviations and acronyms

This glossary uses the following abbreviations and acronyms:

AI	Artificial Intelligence
AI-MD	AI based medical device
AI4H	Artificial Intelligence for health
API	Application Programming Interface
CfTGP	Call for Topic Group Participation
CNN	Convolutional Neural Network
COA	Clinical outcome assessment
CONSORT-AI	Consolidated Standards of Reporting Trials
DEL	Deliverable
DL	Deep Learning
FDA	Food and Drug Administration
FG	Focus Group
FG-AI4H	Focus Group on AI for Health
GA	Genetic Algorithm

GDPR	General Data Protection Regulation
IMDRF	International Medical Device Regulators Forum
IP	Intellectual property
ISO	International Standardization Organization
ITU	International Telecommunication Union
LMIC	Low-and middle-income countries
MDR	Medical Device Regulation
ML	Machine Learning
NGO	Non-Governmental Organization
NN	Neural Networks
SaMD	Software as a Medical Device
SDG	Sustainable Development Goal
TG	Topic Group
TPLC	Total Product Life Cycle
UN	United Nations
WG	Working Group
WHO	World Health Organization

Bibliography

[ITU-T A.1] Recommendation ITU-T A.1, *Working methods for study groups of the ITU Telecommunication Standardization Sector*.
https://www.itu.int/rec/T-REC-A.1

[ITU-T A.7] Recommendation ITU-T A.7, *Focus groups: Establishment and working procedures*.
https://www.itu.int/rec/T-REC-A.7

[ITU-T H.764] Recommendation ITU-T H.764 (2012), *IPTV services enhanced script language*.
https://www.itu.int/rec/T-REC-H.764

[ITU-T X.1252] Recommendation ITU-T X.1252, *Baseline identity management terms and definitions*
https://www.itu.int/rec/T-REC-X.1252/en

[ITU-T Y.3172] Recommendation ITU-T Y.3172, *Architectural framework for machine learning in future networks including IMT-2020*
https://www.itu.int/rec/T-REC-Y.3172/en

[ITU-T FG-DPM D0.1] FG DPM – Focus Group on Data Processing and Management to support IoT and Smart Cities & Communities (2019), *Technical Specification D0.1 - Data Processing and Management for IoT and Smart Cities and Communities: Vocabulary*
http://handle.itu.int/11.1002/pub/813b0843-en

[CONSORT-AI] Liu, X. et al. (2020), *Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension.* Nature Medicine vol. 26 1364–1374.
https://doi.org/10.1136/bmj.m3164

[CTA-2089] Definitions and Characteristics of Artificial Intelligence (ANSI/CTA-2089) (February 2020).
https://shop.cta.tech/products/definitions-and-characteristics-of-artificial-intelligence

[Deep Learning] Ian Goodfellow and Yoshua Bengio and Aaron Courville (2016). Deep Learning. *MIT Press*. p. 326.
https://www.deeplearningbook.org/contents/convnets.html

[EC 2021] European Commission (2021). Regulation of The European Parliament And Of The Council – Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206

[Ekelund 2012] Ekelund, Suzanne (2012*). ROC Curves – What are They and How are They Used*? Point of Care: The Journal of Near-Patient Testing & Technology 11(1). p. 16-21.
https://doi.org/10.1097/POC.0b013e318246a642

[FDA 2019] U.S. Food and Drug Administration (2019), FDA-2019-N-1185, *Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) – Discussion Paper and Request for Feedback.*
https://www.regulations.gov/document/FDA-2019-N-1185-0001

[FDA 2021] U.S. Food and Drug Authority (2021). *Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products*. Draft Guidance for Industry.
https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory

[FDA-NIH BEST] FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource [Internet]. Silver Spring (MD): Food and Drug Administration (US); 2016. Co-published by National Institutes of Health (US), Bethesda (MD)
https://www.ncbi.nlm.nih.gov/books/NBK326791/

[IMDRF SaMD-N12] IMDRF SaMD WG/N12FINAL:2014, *Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations.*
http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf

[IMDRF/SaMD-N41] IMDRF SaMD WG/N41FINAL:2017, *SaMD: Clinical Evaluation (N41).*
http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-170921-samd-n41-clinical-evaluation_1.pdf

[IMDRF/AIMD-N67] IMDRF AIMD WG/N67(PD1):2022. Machine Learning-enabled Medical Devices – A subset of Artificial Intelligence-enabled Medical Devices: Key Terms and Definitions.
https://www.imdrf.org/sites/default/files/2022-05/IMDRF%20AIMD%20WG%20Final%20Document%20N67.pdf

[IMDRF/MDCE-N57] IMDRF MDCE WG/N57FINAL:2019, *Clinical Investigation.*
http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-191010-mdce-n57.pdf

[ISO 5127] ISO 5127:2017, *Information and documentation — Foundation and vocabulary*
https://www.iso.org/standard/59743.html

[ISO 9000] ISO 9001:2015, *Quality management systems – Requirements*.
https://www.iso.org/standard/62085.html

[ISO/IEC 22989] ISO/IEC 22989:2022, *Information technology – Artificial intelligence – Artificial intelligence concepts and terminology*.
https://www.iso.org/standard/74296.html

[ISO/IEC TR 24028] ISO/IEC TR 24028:2020, *Information technology – Artificial intelligence – Overview of trustworthiness in artificial intelligence*.
https://www.iso.org/standard/77608.html

[ISO/IEC 25010] ISO/IEC 25010:2011, *Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models*
https://www.iso.org/standard/35733.html

[ISO/IEC 27037] ISO/IEC 27037:2012, *Information technology – Security techniques – Guidelines for identification, collection, acquisition and preservation of digital evidence*
https://www.iso.org/standard/44381.html

[ISO/IEC DIS 25059] ISO/IEC DIS 25059:2023, *Software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – Quality model for AI* systems
https://www.iso.org/standard/80655.html

[ISO/IEC TS 5723] ISO/IEC TS 5723:2022, *Trustworthiness – Vocabulary*
https://www.iso.org/standard/81608.html

[ISO/IEC TS 6254] ISO/IEC AWI TS 6254, *Information technology – Artificial intelligence – Objectives and approaches for explainability of ML models and AI systems*
https://www.iso.org/standard/82148.html

[ISO/IEC 27036-1] ISO/IEC 27036-1:2021, *Cybersecurity – Supplier relationships – Part 1: Overview and concepts*
https://www.iso.org/standard/82905.html

[ISO/TR 20416] ISO/TR 20416:2020, *Medical devices – Post-market surveillance for manufacturers*.
https://www.iso.org/standard/67942.html

[ISO 13485] ISO 13485:2016, *Medical devices – Quality management systems – Requirements for regulatory purposes*.
https://www.iso.org/standard/59752.html

[Nilsson 1998] Nilsson, N. (1998), *Artificial Intelligence: A new synthesis.* Stanford University. Morgan Kauffmann Publishers, Inc.

[PAS 185] PAS 185:2017, *Smart cities – specification for establishing and implementing a security-minded approach*
https://www.cpni.gov.uk/resources/introduction-pas-1852017

[SPIRIT-AI] Cruz Rivera, S. et al. *Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension*. Nat. Med. 26, 1351–1363 (2020).
https://doi.org/10.1136/bmj.m3210

[UN-SDGs] United Nations Sustainable Development Goals (visited 2021-09-01)
https://sdgs.un.org/goals

[WHO AI-EG] World Health Organization Global (28 June 2021), *WHO Guidance – Ethics and governance of artificial intelligence for health*.
https://www.who.int/publications/i/item/9789240029200

[WHO GHE] World Health Organization Global Network of WHO Collaborating Centres for Bioethics (2015). Global Health Ethics. Key issues. ISBN 978 92 4 154911 0. ISBN (PDF) 978 92 4 069403 3.
https://www.afro.who.int/sites/default/files/2017-06/9789240694033_eng.pdf

[Yerushalmy1947] Yerushalmy J (1947). *Statistical problems in assessing methods of medical diagnosis with special reference to X-ray techniques*. *Public Health Reports*. 62 (2): 1432–39. https://doi.org/10.2307/4586294