IMDRF International Medical Device Regulators Forum

Proposed Document

Title: **Machine Learning-enabled Medical Devices—A subset of Artificial Intelligence-enabled Medical Devices: Key Terms and Definitions**

Authoring Group: IMDRF AIMD Working Group

Date: 16 September 2021

Table of Contents

Preface

The document herein was produced by the International Medical Device Regulators Forum (IMDRF), a voluntary group of medical device regulators from around the world. The document has been subject to consultation throughout its development.

1.0 Introduction

Artificial Intelligence (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 computer algorithms to learn through data, without being explicitly programmed, to perform a task.

Approaches utilizing ML, sometimes colloquially referred to as AI or AI/ML, have been employed in several fields, such as the automotive industry, robotics, medicine, finance, and art. ML has given many sectors an ability to gain new insights from large amounts of data and to support tasks.

There has been accelerated adoption and use of ML-enabled approaches in medical devices. We refer to these medical devices as Machine Learning-enabled Medical Devices, or MLMD. AI systems are typically implemented as software in medical devices or as Software as a Medical Device. MLMD have the potential to transform health care by deriving new and important insights from the vast amount of data generated during all phases of the healthcare process. Examples of applications include earlier disease detection and diagnosis; identification of new observations or patterns on human physiology; development of personalized diagnostics and therapeutics; workflow optimization; and guidance in use of the device with the goal of improving user and patient experience. One of the greatest benefits of MLMD resides in its ability to learn from real-world use and experience to improve its performance.

The purpose of this publication is to establish relevant terms and definitions across the Total Product Life Cycle (TPLC) to promote consistency, support global harmonization efforts, and provide a foundation for the development of future guidelines related to MLMD. Terms referenced herein have either been previously defined in Global Harmonization Task Force (GHTF) documents or by internationally recognized standards on AI., Some terms and definitions have been generated by or are discussed by the IMDRF Artificial Intelligence Medical Device (AIMD) Working Group within this document.

The overarching objective of this effort is to promote consistent expectations and understanding for MLMD, promote patient safety, foster innovation, and encourage access to advances in healthcare technology.

2.0 Scope

This document applies to key terms and definitions relating to Machine Learning-enabled Medical Devices (MLMD).

Note 1 : MLMD are *medical devices*. A product must first meet the definition of a *medical device* before it can be an MLMD.

Note 2 : Most jurisdictions include "accessories to medical devices" in the definition of "medical device". Other jurisdictions define "accessories to medical devices" separately. The definitions and the concepts in this document are intended to apply in both case.

Note 3 : This document does not attempt to define established definitions in the field of computer science; however, it does strive to highlight and clarify conflicting terms and definitions as necessary. This document does not provide guidelines for the development, risk management or evaluation of MLMD.

Note 4 : Terms and definitions that refer technical standards that are under development (e.g., ISO, IEC, IEEE) may be updated upon final publication of those standards.

3.0 References

3.1 IMDRF / GHTF

- IMDRF/SaMD WG/N10 FINAL:2013 Software as a Medical Device (SaMD): Key Definitions
- IMDRF/GRRP WG/N47:2018 Essential Principles of Safety and Performance of Medical Devices and IVD Medical Devices(3.0 Definitions)

3.2 Standards

The standards below were consulted in the writing of this document and may be useful in meeting the key definition of MLMD discussed herein. This list is not intended as a required or complete list of standards that can be used to meet the key definition of MLMD.

- ISO/IEC DIS 22989 Information technology — Artificial intelligence — Artificial Intelligence Concepts and Terminology

3.3 Other Documents

- AAMI, BSI, Turpin, R., Hoefer, E., Lewelling, J., & Baird, P. (2020). *Machine Learning AI in Medical Devices: Adapting Regulatory Frameworks and Standards to Ensure Safety and Performance*. AAMI/BSI Initiative on Artificial Intelligence. *https://www.bsigroup.com/en-US/medical-devices/resources/Whitepapers-and-articles/machine-learning-ai-in-medical-devices/*

- Kohavi, R., & Provost, F. (Eds.). (n.d.). Glossary of Terms: Special Issue on Applications of Machine Learning and the Knowledge Discovery Process. *https://ai.stanford.edu/~ronnyk/glossary.html*
- Kan A. (2017). Machine learning applications in cell image analysis. *Immunology and Cell Biology*, *95*(6), 525–530. *https://doi.org/10.1038/icb.2017.16*

4.0 General Overview of Artificial Intelligence and Machine Learning Concepts

AI systems are able to perform tasks such as visual perception, speech recognition, decision-making, and translation between languages – by using expert systems (based on rules like decision trees), machine learning or deep learning.

Some AI systems demonstrate a degree of autonomy (capacity to perform tasks in a complex environment without constant guidance/input from a user) and a capacity for adaptability (ability to learn from experience and thereby change performance.)

ML is a computing system that contains a model or algorithm generated through a computer learning patterns from data, including classification, inference, or matching previous patterns, predicting future outputs, etc. ML has been considered as a subset of AI that gives computers the ability to learn without being explicitly programmed[1].

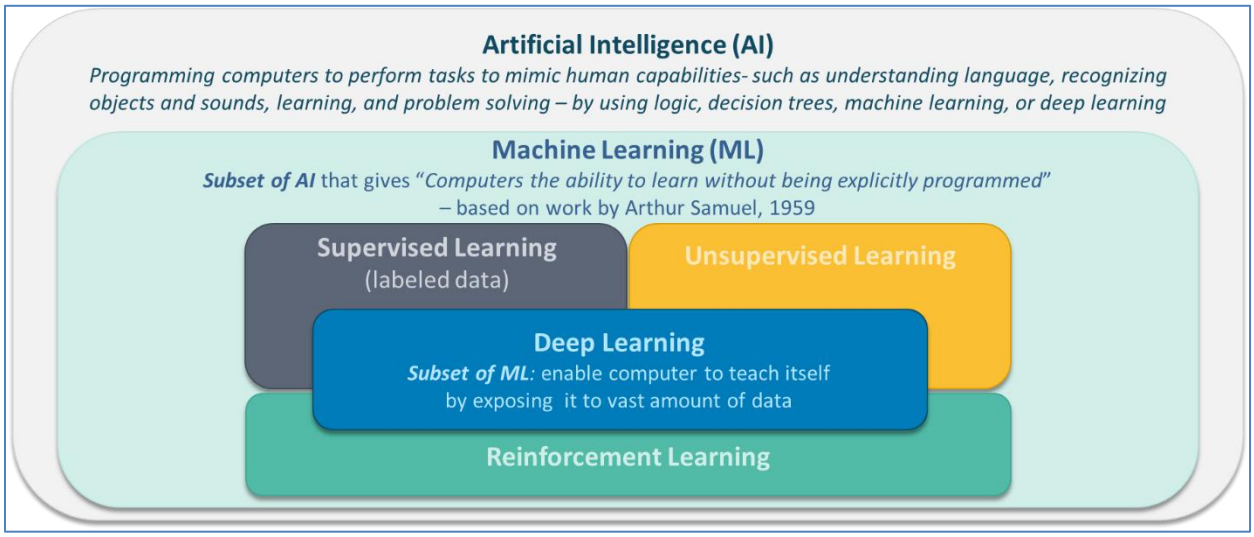

Figure 1 Overview of AI and ML Concepts

ISO/IEC's draft international standard for AI, 22989, defines and discusses ML in terms of being an ML model parameter optimisation process for the purpose of the ML model's behaviour reflecting the data or experience.

There are several different types of ML methods, as well as different algorithms. For example, some applications may use Supervised Learning, others may use Unsupervised or Semi-Supervised Learning (Section 6.0). Different types of algorithms include neural networks (e.g., feed forward neural network, recurrent neural network, convolutional neural network, etc.)

[1] A.L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers." IBM Journal 1(3), 210–229 (1959)

Bayesian networks, decision trees, support vector machine, among others. The learning process itself may be an iterative process of trial and error, also known as Reinforcement Learning.

Note : Within this document, the term ML algorithm is used to represent a software procedure developed using ML, and consisting of mathematics and logic, that can process data. The term ML model is used here to represent the relationship or function that is the result of Training an ML algorithm with data.

The following sections provide key definitions that are relevant to ML when used in medical devices (Section 5.0), definitions from technical standards (Section 6.0), followed by a discussion of common ML terms (Section 7.0).

5.0 Key Definitions

5.1 Machine Learning-enabled Medical Device (MLMD)

A medical device that uses machine learning, in part or in whole, to achieve its intended medical purpose.

5.2 IMDRF Terms

Medical Device: Any instrument, apparatus, implement, machine, appliance, implant, reagent for in vitro use, software, material or other similar or related article, intended by the manufacturer to be used, alone or in combination, for human beings, for one or more of the specific medical purpose(s) of:

- diagnosis, prevention, monitoring, treatment or alleviation of disease,
- diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury,
- investigation, replacement, modification, or support of the anatomy, or of a physiological process,
- supporting or sustaining life,
- control of conception,
- cleaning, disinfection or sterilization of medical devices,
- providing information by means of in vitro examination of specimens derived from the human body;

and does not achieve its primary intended action by pharmacological, immunological, or metabolic means, in or on the human body, but which may be assisted in its intended function by such means.

Note 1 : Products which may be considered to be medical devices in some jurisdictions but not in others include:

- disinfection substances,
- aids for persons with disabilities,
- devices incorporating animal and/or human tissues,
- devices for in-vitro fertilization or assisted reproduction technologies.

Note 2 : For clarification purposes, in certain regulatory jurisdictions, devices for cosmetic/aesthetic purposes are also considered medical devices.

Note 3 : For clarification purposes, in certain regulatory jurisdictions, the commerce of devices incorporating human tissues is not allowed.

Editorial issue has been corrected from IMDRF/GRRP WG/N47:2018.

6.0 Definitions/Reference Definitions/ Technical Standards Definitions

6.1 Bias

Systematic difference in treatment of certain objects, people, or groups in comparison to others. (ISO/IEC DIS 22989)

Note : Bias is used in both data science and in legal discussions. When used in data science, bias is the tendency of a statistic to overestimate or underestimate a parameter. From a legal point of view, however, bias is any prejudiced or partial personal or social perception of a person or group. For the purposes of this document, bias is a data science term, and not a legal one. Bias can be introduced into study design, conduct or analysis. Sources of bias include selection bias (of study sample), operational bias, and analyses that do not account for missing data.

6.2 Continuous Learning

Training that leads to change of an MLMD with each exposure to data that takes place on an ongoing basis during the operation phase of the MLMD life cycle. (Modified from ISO/IEC DIS 22989)

Note : Batch Learning is a training that leads to the change of an MLMD that involves discrete updates based on defined sets of data that take place at distinct points prior to or during the operation phase of the MLMD life cycle.

6.3 Reference Standard

An objectively determined benchmark that is used as the expected result for comparison, assessment, training, etc. (e.g., ground truth, gold standard).

6.4 Reinforcement Learning

Machine learning utilizing a reward function to optimize either a policy function or a value function by sequential interaction with an environment. (ISO/IEC DIS 22989)

Note 1 to entry: Policy functions and value functions express a strategy that is learned by the environment.

Note 2 to entry: The environment can be any stateful model.

6.5 Reliability

Property of consistent intended behavior and results. (ISO/IEC DIS 22989)

6.6 Semi-Supervised Machine Learning

Machine learning that makes use of both labelled and unlabelled data during training. (ISO/IEC DIS 22989)

Note 1 : Descriptive information can be broader than just labelling. Annotation is the process of attaching descriptive information to data, such as metadata, labels, or anchors. The data itself is unchanged in the annotation process.[2]

Note 2 : Additional information about this term can be found in Section 7.4

6.7 Supervised Machine Learning

Machine learning that makes use of labelled data during training. (ISO/IEC DIS 22989)

Note 1 : Descriptive information can be broader than just labelling. Annotation is the process of attaching descriptive information to data, such as metadata, labels, or anchors. The data itself is unchanged in the annotation process.[2]

Note 2 : Additional information about this term can be found in Section 7.4

6.8 Test Dataset

A subset of the data that is never shown to the ML model during training, used to verify what the model has learned. (Modified from ISO/IEC DIS 22989)

6.9 Training

Process intended to establish or to improve the parameters of a machine learning model, based on a machine learning algorithm, by using training data. (Modified from ISO/IEC DIS 22989)

6.10 Training Dataset

Subset of input data samples used to train a machine learning model. (ISO/IEC DIS 22989)

6.11 Unsupervised Machine Learning

Machine learning that makes use of unlabelled data during training. (ISO/IEC DIS 22989)

Note 1 : Descriptive information can be broader than just labelling. Annotation is the process of attaching descriptive information to data, such as metadata, labels, or anchors. The data itself is unchanged in the annotation process.[2]

Note 2 : Additional information about this term can be found in Section 7.4

[2] ISO/IEC DIS 22989 Information technology — Artificial intelligence — Artificial Intelligence Concepts and Terminology

7.0 Discussion

The following sub-sections contain discussions of concepts that warranted more detail than a concise definition. In particular, the aspects of MLMD changes, supervised and unsupervised learning, and validation are discussed.

7.1 Aspects of MLMD Changes

MLMD offer unique benefits, flexibility, and challenges related to their capacity for change. The transparent communication of the various aspects of these changes is important to the safety, performance, and effectiveness of MLMD.

The examples outlined in this discussion are not exhaustive and the relevant information may expand over time. It is important to note that changes , such as software patches, operating system updates, cybersecurity improvements, etc., can impact both MLMD and non-MLMD medical devices and, although important, these changes are not within the scope of this discussion.

There are a number of unique changes related to MLMD, including changes to the ML model or to the environment of use relative to the ML training data. The following discussion highlights these important aspects in two sections, MLMD Changes and MLMD Environmental Changes.

7.1.1 Changes to MLMD

A change to the device could include a modification to the machine learning model, algorithm, weights, or parameters. MLMD is in a locked state when changes are not permitted. Aspects that describe these changes include the cause, effect, trigger, domain, timing, and effectuation. These attributes describe what changes, as well as why, where, when, and how the MLMD change occurs.

Note : The word "locked" has been used by the community in a number of different ways. Some have defined a "locked device" as one that has been developed using ML methods but for which the developer does not have an intention of modifying at the present time. Others have used the term "locked device" as any device that does not perform "continuous learning." When using the word "locked" it is important to provide clarifying language around its use to communicate how it is being used.

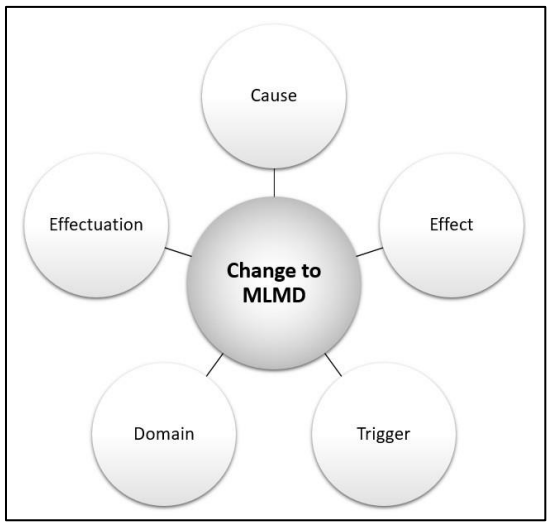

Figure 2 Aspects of MLMD Changes

The cause refers to the source of the change to the MLMD, for example, re-training with new or appended data or new training methods, algorithm/model, tuning, etc.

The effect refers to the resulting change to the MLMD, which can include amended intended use/indications for use; modified performance, changes in inputs, outputs, etc.

The trigger refers to the event that prompts or instigates the change to the MLMD, which can include performance thresholds, training data batch-size thresholds, exposure to new data/experiences, scheduled time intervals, MLMD environmental changes, user feedback, etc.

The domain refers to the scope or applicable extent of the change to the MLMD, which can be categorized as either homogeneous or heterogeneous. A homogeneous change is a uniform change that occurs universally (sometimes referred to as a global adaptation, note that global does not

denote around-the-world). Heterogeneous changes are non-uniform changes that can be specific to one clinic, region, demographic, etc. (sometimes referred to as local adaptations).[3]

The effectuation refers to where the mechanism for change implementation resides, which can either be external (i.e., updated by the developer or user) or internal (i.e., updated by a change-control-algorithm within the device).

7.1.2 Changes to MLMD Environment

An MLMD environmental change is a modification to the setting of the MLMD relative to the ML development data. Aspects that describe an MLMD environmental change include the cause, effect, and domain.

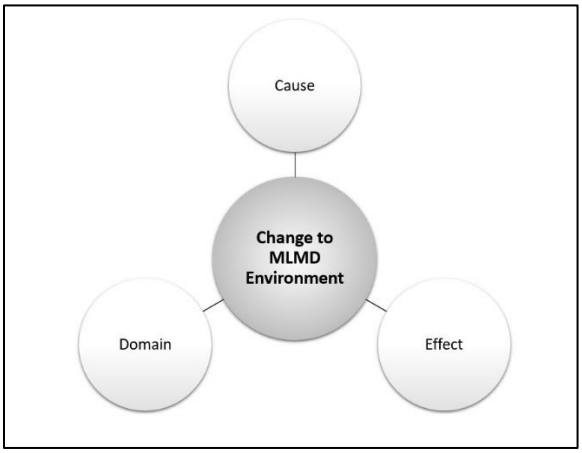

Figure 3 Aspects of MLMD Environmental Changes

The cause of an MLMD environmental change refers to the source of the change relative to the development environment. Examples of such causes include changes to the format or quality of the MLMD inputs (e.g., changes to third party image processing, incidents of adversarial machine learning); changes in the patient population (e.g., demographic shift); changes in clinical practice (e.g., earlier interventions that mask features used by the model for classification), etc.

The effect of an MLMD environmental change can involve deteriorated or improved performance, effectiveness, or safety.

The domain of an MLMD environmental change refers to the scope or applicable extent of the change, which can be categorized as either homogeneous or heterogeneous. Heterogeneous changes are non-uniform changes that can be specific to one clinic, region, demographic, etc. (sometimes referred to as local changes). Homogeneous changes are changes that occur uniformly (universally, globally) over some groups or settings/context. Note that global does not denote around-the-world.

[3] "Introduction to Online Machine Learning: Simplified", https://www.analyticsvidhya.com/blog/2015/01/introduction-online-machine-learning-simplified-2/

7.2 Supervised / Unsupervised / Semi-Supervised Learning

Supervised and Unsupervised Machine Learning are two methods that are commonly used to train machine learning algorithms, but they are not the only methods available. The terms "supervised" and "unsupervised" in a machine learning context refer to the training methods, and specifically whether labelled or unlabelled data are used. Supervised Machine Learning utilizes labelled data during Training to learn the relationship between independent attributes and a designated dependent attribute (the label). In other words, supervised learning is a task to learn a mapping from input to output values, where the correct output values are known (labelled training data). Most induction algorithms are developed through supervised learning. Unsupervised Machine Learning utilizes unlabelled data during Training to group data without a pre-specified dependent attribute. In other words, unsupervised learning is the ability to find patterns from input values, where the output values are unknown. Examples of unsupervised learning include some types of algorithms that perform clustering or dimensionality reduction.

Machine learning systems can use a mix of supervised and unsupervised learning (sometimes referred to as semi-supervised learning), as well as other learning methods such as Reinforcement Learning.

The terms "Supervised Machine Learning " and "Unsupervised Machine Learning " are often misunderstood. When used in a machine learning context, "supervised" or "unsupervised" does not refer to the presence or absence of a human supervisor overseeing the software. "Supervised" or "unsupervised" does not refer to the role that the software plays in a clinical environment, i.e., it does not describe the level of "autonomy" in practice. "Supervised" or "unsupervised" also does not refer to whether the software updates itself in a self-effectuating update process, i.e., whether it performs its own updates or adaptations.

7.3 Validation

The term validation has been used to represent different concepts within the fields of medical device development and machine learning algorithm development.

Validation within the context of medical device development has been defined as follows:

> *Validation means confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use can be consistently fulfilled.*[4]

The term validation has also been used within the field of machine learning to refer to either data curation (sometimes referred to as data validation) or model tuning (sometimes referred to as validation[5]).

[4]Design Control Guidance for Medical Device Manufacturers (GHTF.SG3.N99-9)

[5]Ripley, B. (1996). Glossary. In Pattern Recognition and Neural Networks (pp. 347-354). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511812651.013

Data curation and model tuning can occur throughout the product lifecycle. Data curation refers to the selection, management and assessment of the quality attributes of data sets. Model tuning is a particular phase of model development during which ML model hyper-parameters are tuned; this optional tuning phase can be combined with the Training phase to optimize the ML model selection.

MLMD manufacturers, regulators, and users should be aware of the conflicting interpretations of the term validation and ensure that communication regarding the development phases and the associated datasets is clear to avoid confusion between data validation, model tuning, and medical device validation. It is recommended that the use of the term “validation” be accompanied by the context when referring to model tuning, data curation, and the associated datasets. Alternatively, the use of the term validation that refers to the training and tuning process may be avoided in the context of medical device development.