

**NEMA MITA JSC 51500-2023**

*Machine Learning Algorithms:  
Dataset Management Best Practices in Medical Imaging*

*Published by:*

**National Electrical Manufacturers Association**  
1300 North 7th Street, Suite 900  
Rosslyn, Virginia 22209

[www.nema.org](http://www.nema.org)

© 2023 by the National Electrical Manufacturers Association. All rights including translation into other languages, reserved under the Universal Copyright Convention, the Berne Convention for the Protection of Literary and Artistic Works, and the International and Pan American Copyright Conventions.

## NOTICE AND DISCLAIMER

The information in this publication was considered technically sound by the consensus of persons engaged in the development and approval of the document at the time it was developed. Consensus does not necessarily mean that there is unanimous agreement among every person participating in the development of this document.

The National Electrical Manufacturers Association (NEMA) standards and guideline publications, of which the document contained herein is one, are developed through a voluntary consensus standards development process. This process brings together volunteers and/or seeks out the views of persons who have an interest in the topic covered by this publication. While NEMA administers the process and establishes rules to promote fairness in the development of consensus, it does not write the document and it does not independently test, evaluate, or verify the accuracy or completeness of any information or the soundness of any judgments contained in its standards and guideline publications.

NEMA disclaims liability for any personal injury, property, or other damages of any nature whatsoever, whether special, indirect, consequential, or compensatory, directly or indirectly resulting from the publication, use of, application, or reliance on this document. NEMA disclaims and makes no guaranty or warranty, express or implied, as to the accuracy or completeness of any information published herein, and disclaims and makes no warranty that the information in this document will fulfill any of your particular purposes or needs. NEMA does not undertake to guarantee the performance of any individual manufacturer or seller's products or services by virtue of this standard or guide.

In publishing and making this document available, NEMA is not undertaking to render professional or other services for or on behalf of any person or entity, nor is NEMA undertaking to perform any duty owed by any person or entity to someone else. Anyone using this document should rely on his or her own independent judgment or, as appropriate, seek the advice of a competent professional in determining the exercise of reasonable care in any given circumstances. Information on other standards on the topic covered by this publication may be available from other sources, which the user may wish to consult for additional views or information not covered by this publication.

NEMA has no power, nor does it undertake to police or enforce compliance with the contents of this document. NEMA does not certify, test, or inspect products, designs, or installations for safety or health purposes. Any certification or other statement of compliance with any health or safety-related information in this document shall not be attributable to NEMA and is solely the responsibility of the certifier or maker of the statement.

## Contents

Foreword .....	ii
Introduction.....	iii
1 Scope .....	1
2 Terms and Definitions .....	1
3 Development Planning.....	2
3.1 Algorithm Purpose .....	3
3.2 Scope of Use .....	3
3.3 Context of Use .....	3
3.4 Data Quality Criteria .....	3
4 Data Collection.....	4
5 Data Annotation .....	5
6 Data Quality Assessment.....	5
6.1 Data Appropriateness .....	5
6.2 Reference Standard .....	5
7 Dataset Partitioning.....	6
7.1 Dataset Sizes.....	6
7.2 Dataset Composition .....	7
7.3 Dataset Segregation.....	7
8 Dataset Maintenance.....	7
9 Conclusion .....	7
Annex A (Informative) Synthetic Data.....	9

## Figures

Figure 1 – Dataset Diagram.....	2
---------------------------------	---

## Foreword

NEMA technical publications are subject to periodic review. They are revised frequently to reflect user input and to meet changing conditions and technical progress. Proposed revisions to this technical publication should be submitted to:

NEMA Technical and Industry Affairs Department  
National Electrical Manufacturers Association  
1300 North 17<sup>th</sup> Street, Suite 900  
Rosslyn, VA 22209

Note: The user's attention is called to the possibility that compliance with this standard could require use of an invention covered by patent rights.

By publication of this standard, no position is taken with respect to the validity of any such claim(s) or of any patent rights in connection therewith. If a patent holder has filed a statement of willingness to grant a license under these rights on reasonable and non-discriminatory terms and conditions, applicants desiring to obtain such a license, then details may be obtained from the Secretary.

## Introduction

This document captures present-day best practices for data management related to machine learning (ML) algorithm development within the medical imaging device industry. These best practices are presented as a framework that developers can reference as a starting point for product risk analysis or for process development as a quality management system component when using machine learning algorithms to generate output. Practices described and any related examples do not prescribe what is necessary in algorithm development, but instead serve to illustrate the unique considerations machine learning algorithms may involve during the product development process.

This document was also developed in response to a growing interest from other healthcare sector stakeholders (such as clinicians and regulators) for additional clarity on the machine learning algorithm development process for medical imaging devices. Manufacturers are committed to the safety and efficacy of all imaging products, and support policies and standards that deliver improved patient outcomes and support public health. However, as noted above, the best practices enumerated in this document in no way imply any requirement to provide information beyond that necessary for safe and effective use.

The framework is intended to help developers create understandable data quality statements or other documentation that communicates how data quality was addressed within ML-based development processes. The topics covered in this document may be useful considerations for what information could be provided by the developer to the end user or other stakeholders. This document does not describe how to meet regulatory requirements, such as labeling requirements.

**<This page is intentionally left blank.>**

Currently in preview, click buy full version

## 1 Scope

This document addresses considerations for evaluating data for quality, fitness, and usability to achieve the intended use<sup>1</sup>, determined by the manufacturer, for an ML imaging device. To that end, the best practices discussed in this document focus on the development and management of good quality datasets.

This document divides those practices into six sections:

1. Development Planning
2. Data Collection
3. Data Annotation
4. Data Quality Assessment
5. Dataset Partitioning
6. Dataset Maintenance

This document considers data quality issues that may directly impact ML algorithm performance in a medical imaging device. Issues out of scope for this document include:

- quality considerations for non-AI/ML device components;
- legal considerations, such as data privacy laws and regulations;
- medical ethics considerations;
- risk evaluation for automated processes;
- ecological and social impact considerations;
- general medical device development processes and quality management systems
- cybersecurity concerns, such as detecting dataset corruption, preventing tampering, etc.;
- data availability, patient consent concerns (whether the data can be used or not);
- regulatory considerations and post-market data practices

These topics are important and may be addressed in future publications.

## 2 Terms and Definitions

The terms “test,” “train,” “evaluate,” and “validate” are often used in different ways by data scientists, algorithm developers, medical device manufacturers, medical device regulators, and other stakeholders. Such terms are defined below to clarify their use and meaning within this paper.

**machine learning (ML) model:** A mathematical construct that generates an inference or prediction based on new input data and is the result of an ML training algorithm learning from data<sup>2</sup>.

**data record:** A collection of related data elements and associated metadata. A data record typically corresponds to the input available to a single execution of a machine learning algorithm, e.g., the data associated with a single imaging study being analyzed for pneumonia fed into the algorithm and a result produced.

**total dataset:** The complete set of collected data for the identified project, i.e., the development and evaluation datasets.

<sup>1</sup> IMDRF/GRRP WG/N52 FINAL:2019

<sup>2</sup> IMDRF/AIMD WG/N67, "Machine Learning-enabled Medical Devices: Key Terms and Definitions" (<https://www.imdrf.org/documents/machine-learning-enabled-medical-devices-key-terms-and-definitions>)