# MACHINE LEARNING AI IN MEDICAL DEVICES

## Adapting Regulatory Frameworks and Standards to Ensure Safety and Performance


# ABOUT

## AAMI

AAMI is a nonprofit organization founded in 1967. It is a diverse community of approximately 9,000 professionals united by an important mission—the development, management, and use of safe and effective health technology.

AAMI is a primary source of consensus standards, both national (US) and international, for medical device industry, as well as practical information, support, and guidance for healthcare technology and sterilization professionals.

## BSI

BSI is a global thought leader in the development of standards of best practice for business and industry. Formed in 1901, BSI was the world's first National Standards Body (NSB) and a founding member of the International Organization for Standardization (ISO). Over a century later, BSI is focused on business improvement across the globe, working with experts in all sectors of the economy to develop codes, guidance and specifications that will accelerate innovation, increase productivity and support growth. Renowned as the originator of many of the world's best-known business standards, BSI's activity spans multiple sectors including aerospace, automotive, built environment, food, healthcare, and ICT.

Over 95% of BSI's work is on international and European standards. In its role as the UK National Standards Body, BSI represents UK economic and social interests across the international standards organizations ISO, IEC, CEN, CENELEC and ETSI, providing the infrastructure for over 11,000 experts to work on international, European, national, and PAS standards development in their chosen fields.

## AAMI/BSI INITIATIVE ON AI

The AAMI/BSI Initiative on Artificial Intelligence (AI) in medical technology is an effort by AAMI and BSI to explore the ways that AI and, in particular, machine learning pose unique challenges to the current body of standards and regulations governing medical devices and related technologies. Also to determine what additional guidance or standards might be needed to promote the safety and effectiveness of medical AI technologies. Two stakeholder workshops to explore the issue were held in the fall of 2018 and resulted in the publication of a first whitepaper, *The emergence of artificial intelligence and machine learning algorithms in healthcare: Recommendations to support governance and regulation.*

This second whitepaper builds on that initial work and was informed by two additional workshops held in Arlington, VA, USA, and in London, United Kingdom in 2019.

### PRIMARY AUTHORS

Rob Turpin – BSI
Emily Hoefer – AAMI
Joe Lewelling – AAMI
Pat Baird – Philips

 

# ACKNOWLEDGEMENTS

The following individuals participated in the 2019 workshops or otherwise provided input to the authors. Contributions of their time and expertise are greatly appreciated.

**LONDON WORKSHOP PARTICIPANTS:**

Neil Ebenezer, UK Department for International Trade
David Grainger, UK Medicines and Healthcare products Regulatory Agency (MHRA)
Stuart Harrison, ETHOS
Joanne Holden, National Institute for Health and Care Excellence (NICE)
Xiaoxuan Liu, University Hospitals Birmingham NHS Foundation Trust
Charles Lowe, Digital Health and Care Alliance (DHACA)
Sheena MacPherson, Miotify Ltd.
Jacqueline Mathews, NIHR Clinical Research Network
Damien McPhillips, Boston Scientific
Deborah Morrison, National Institute for Health and Care Excellence (NICE)
Alberto Rizzoli, V7 Ltd.
Danny Ruta, Guy's and St. Thomas' NHS Foundation Trust
Antoine Saillant, Feedback Medical Ltd.
Richard Scott, IEC/TC 62
Haris Shuaib, Guy's and St. Thomas' NHS Foundation Trust

**ARLINGTON, VA WORKSHOP PARTICIPANTS:**

Lisa Carnahan, National Institute of Standards and Technology (NIST)
Melanie Darovitz, Kaiser Foundation Health Plan/Hospitals
Kathryn Drzewiecki, US Food and Drug Administration/Center for Devices and Radiological Health (FDA/CDRH)
Marc Edgar, General Electric Healthcare – Digital
Jesse Ehrenfeld, American Medical Association
Lars Lynne Hansen, Novo Nordisk
Lacey Harbour, Harbour Regs LLC
Zack Hornberger, Medical Imaging & Technology Alliance (MITA)
Patrick Jones, Philips
Cameron Loper, MPR Associates Inc
Christina Silcox, Duke-Margolis Center for Health Policy
Andrew Southerland, Departments of Neurology and Public Health Sciences University of Virginia
Scott Thiel, Navigant Consulting Inc
Sylvia Trujillo, American Medical Association (Formerly) and Compassion & Choices (currently)
Jamie Wolszon, Advanced Medical Technology Association (AdvaMed)
Krista Woodley, Johnson & Johnson

*NOTE—Participation in the workshops or in the review of this whitepaper by any individual, including government agency representatives, does not constitute endorsement or approval of its contents by those individuals or agencies.*

# EXECUTIVE SUMMARY

This paper examines how AI is different from traditional medical devices and medical software, explores the implications of those differences, and discusses the controls necessary to ensure AI in healthcare is safe and effective. Because these differences will not be the same for the full range of systems, it is important to identify what aspects of AI are of concern.

**BSI AND AAMI, IN CONSULTATION WITH KEY STAKEHOLDERS, RECOMMEND:**

1. developing, in collaboration with International Medical Device Regulators Forum (IMDRF) and other regulatory bodies, standardized critical terminology and a taxonomy for medical AI that can inform future national and regional regulatory approaches to the technology (Clause 6);
2. that IMDRF establish an AI working group to address issues around AI in healthcare and to prepare needed guidance and good regulatory practices (Clause 11);
3. mapping AI-applicable international regulatory standards (where such exist) to the October 2018 IMDRF Essential Principles and identifying gaps where additional new standards or guidance are needed (Clause 12);
4. developing guidance on factors affecting data quality in regard to AI as a medical technology (Clause 13);
5. establishing a common set of criteria for the deployment of AI in healthcare systems that could be used as an evaluation protocol by multiple stakeholders, covering organizational management, professional conduct, research and ethics, evidence-based practice, and data governance (Clause 14);
6. developing risk management guidance to assist in applying ISO 14971 to AI as a medical technology (Clause 15); and
7. developing guidance on factors to consider in the validation of AI systems and on the use of non-traditional approaches, such as excellence assessments, to demonstrate a reasonable assurance of product quality (Clause 16).

 

# MACHINE LEARNING AI IN MEDICAL DEVICES: ADAPTING REGULATORY FRAMEWORKS AND STANDARDS TO ENSURE SAFETY AND PERFORMANCE

"A robot may not injure a human being or, through inaction, allow a human being to come to harm." *—The First Rule of Robotics, from I Robot, by Isaac Asimov*

"Practice two things in your dealings with disease: either help or do not harm the patient." *—Epidemics, Book I, of the Hippocratic school, attributed to Thomas Inman*

## 1. INTRODUCTION

Artificial Intelligence (AI) promises to revolutionize the practice of medicine by making healthcare more accessible, more efficient, and even more effective. The concept of AI itself, however, is ambiguous if not controversial. AI has been broadly defined as "the capability of a machine to imitate intelligent human behavior".[1] A narrower and more complex definition applies the term to systems that "display intelligent behavior by analyzing their environment and taking actions—with some degree of autonomy—to achieve specific goals…"[2] In the first definition, the systems are merely mimicking human intelligence, but in the second they exhibit a degree of reasoning or cognition—of "thinking" in some sense of the word.

Whatever promise it holds, AI, like any new healthcare technology, can present challenges to existing methods for ensuring safety and performance. The safety and effectiveness[3] of medical devices entering the market today are governed by regulations and private-sector consensus standards. These controls (standards and regulations) were developed alongside current technologies and are based on an extensive, shared understanding of how and how well they work. With an emergent technology like AI, real-world experience is limited, which can hinder our ability to fully assess its effectiveness. Similarly, a lack of real-world experience with AI limits our understanding of its associated risks. AI-related risks are harder to quantify and mitigate; there may be unforeseeable and unpredictable hazards arising from the unique nature or function of AI.

This paper examines how AI is different from other medical devices and medical software, explores the implications of those differences, and discusses the controls necessary to assure AI in healthcare is safe and effective. As these differences will not be the same for the full range of systems, it is important to identify what aspects of AI are of concern.

## 2. MACHINE-LEARNING AI—WHAT MAKES IT DIFFERENT?

Systems that simply imitate humans are not new to the medical field, of course. Since the advent of commercial transistors in the 1960s, computational medical devices have increasingly mimicked human behavior and actions. Automatic blood pressure monitors (sphygmomanometers) imitate the actions of trained clinicians in detecting and reporting the Korotkoff sounds that signify systolic and diastolic blood pressures. Portable defibrillators evaluate heart waveforms to determine when defibrillation is necessary and can then act to deliver the needed defibrillation.

Devices like these, by supplementing or in some instances replacing direct clinician involvement, have already expanded the availability of care outside of healthcare facilities, to homes and workplaces, as well as to areas and regions where trained

*1 Definition from the Merriam Webster Dictionary.*

*2 Communication from the European Commission to the European Parliament, the European Council, the Council Europe and Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe, Brussels, 25.4.2018 COM (201) 237 Final.*

*3 Within this whitepaper the terms "effectiveness" and "performance" with respect to medical devices are used interchangeably in accordance with the definition of "effectiveness" provided in the U.S. Code of Federal Regulations. ("21 CFR 860.7). See A.5*

clinicians are rare or absent. Such technologies, however, do not act independently of human reasoning, but instead utilize previously validated clinical protocols to diagnose medical conditions or deliver therapy. They do not "think" for themselves in the sense of understanding, making judgements, or solving problems[4]; rather, they are static *rules-based systems,*[5] programmed to produce specific outputs based on the values of received inputs.

While such systems can be very sophisticated, the rules they employ are static—they are not created or modified by the systems. Their algorithms are developed based on documented and approved clinical research and then validated to produce expected (i.e., predictable) results. In this aspect, rules-based AI systems, other than their complexity, do not differ substantially from computational and electronic medical devices that have been in use since the 1960s.

There are other types of AI that utilize large data sets and complex statistical methodologies to discover new relationships between inputs, actions, and outcomes. These *data-driven or machine learning systems*[6] are not explicitly programmed to provide pre-determined outputs, but are heuristic, with the ability to learn and make judgements. In short, machine learning AI systems, unlike simple rules-based systems, are cognitive in some sense and can modify their outputs accordingly. For the purposes of this whitepaper, we have separated data-driven/machine learning AI into two groups—locked models that are unable to change without external intervention, and continuous learning (or adaptive models) that modify outputs automatically in real-time (see Figure 1). In reality, there are likely to be several levels of change control for AI—from traditional concepts that are already known, to accelerated concepts that may need additional levels of control.

The more sophisticated of these data-driven systems (i.e., super-intelligent AI) can surpass human cognition in their ability to process enormous and complicated data sets and engage in higher levels of abstraction. Utilizing multiple layers of statistical analysis and deep learning/neural networks, these systems act as *black boxes*[7] producing protocols and algorithms for diagnosis or therapy that are not readily understandable by clinicians or explicable to patients.

*4 One definition of "Think" in the Cambridge Dictionary is "to use one's mind to understand."*

*5 Daniels, et al, Current State and Near-Term Priorities for AI-Enabled Diagnostic Support Software in Health Care (White Paper), Duke Margolis Center for Health Policy, 2019, p. 10*

*6 Ibid. For the purposes of this paper, the terms data-driven, and machine-learning are synonymous, as are the terms continuous learning and adaptive models.*

*7 The metaphor of "black box" is used widely and with different connotations, but with respect to AI, we are not simply talking about a lack of visibility with respect to mechanisms or calculations, but also to the inscrutability of the basic rationale for performance.*

| Rules-Based AI Systems | Data-Driven/Machine-Learning AI Systems | |
|---|---|---|
| | **Locked Machine Learning Models** | **Continuous Learning/Adaptive Models** |
| • Mimic human behavior-making decisions by applying static rules to arrive at predicable decisions. | • Neither the internal algorithms nor system outputs change automatically. | • Utilize newly received data to test assumptions that underlie their operation in real-world use. |
| • Often visualized as a decision tree. | • Further machine learning can be implemented through external approval, or in a stepwise manner. | • Programed to automatically modify internal algorithms and update external outputs in response to improvements being identified. |
| • May be originally developed based on a set of rules provided by human experts or can be based on training data. | | |
| • The logic used to make decisions is usually clear and reproducible. | | |

**Figure 1. Rules-Based System versus Data-Driven Artificial Intelligence Systems**

Data-driven machine learning AI systems can be further divided into *locked models and continuous learning models:*

- *Locked models*[8] employ algorithms that are developed using training data and machine learning, which are then fixed so neither the internal algorithms nor system outputs change automatically (though changes can be accommodated in a stepwise manner).
- *Continuous learning models* (or adaptive models)[9] utilize newly received data to test the assumptions that underlie their operations in real-world use and, when potential improvements are identified, the systems are programed to automatically modify internal algorithms and update external outputs.

The special characteristics of machine learning and deep-learning AI systems differentiate them from rules-based systems and more traditional medical devices in specific ways. First, they learn—these systems not only treat patients, but are capable of assessing the results of treatment both for individuals and across populations, as well as making predictions about improving treatment to achieve better patient outcomes. Second, they are capable of autonomy—some of these systems have the potential to change (and presumably improve) processes and outputs, without direct clinical oversite or traditional validation. Third, because of their sophisticated computational abilities, the predictions developed by these systems may, to some degree, be inexplicable to patients and clinicians. Combined, these characteristics blur the essential nature of the devices themselves, changing them from being simply tools used under the direction of clinicians to systems capable of making autonomous clinical judgements about treatment.

## 3. COMPETENCE, TRUST, AND RELIABILITY

"Never be afraid to trust an unknown future to a known God." *—Corrie ten Boom*

"[W]ith artificial intelligence we're summoning the demon." *—Elon Musk*

For more than 2,000 years, medicine has embraced the ethos of Hippocrates to "first, do no harm." A corollary of that is any action taken to treat a disease, illness, or injury should alleviate the patient's condition in some way—it must be effective. In modern times, ensuring the effectiveness of treatments has relied upon the use of the scientific method. Treatments and procedures must be scientifically established, must be supported by clinical evidence, and should be understood and explainable. Clinicians, who oversee these treatments and perform these procedures must show possession of high levels of scientific knowledge and technical skill and be licensed or accredited. In short, practitioners must demonstrate competence before they are permitted to practice medicine.

Trust[10] with respect to medical devices is different. Traditional rules-based medical devices do not practice medicine, but rather perform automated pre-programmed tasks. For medical devices and technology, acceptable adherence to the scientific method starts with using established scientific principles in device design, followed by conformance to consensus standards that require manufacturers to prove the effectiveness of their products through clinical investigations and empirical evidence, as well as compliance with governmental regulations. Those standards and regulations require that safety be demonstrated through testing and risk management. They also require manufacturers to employ various practices[11] in a quality management system that assure any substantive change to a product's design, materials, manufacture, or function is similarly supported by clinical or empirical evidence. In other words, trust in medical technology is established not by

*8 The term "locked" with respect to AI has been defined as "a function/model that was developed through data-based AI methods, but does not update itself in real time (although supplemental updates can be made to the software on a regular basis)." [Source: Duke, Current State and Near-Term Priorities for AI-Enabled Diagnostic Support Software in Health Care]. A "locked" data-driven algorithm, even if externally validated, is not a rules-based algorithm, because that locked AI algorithm is not based on current, rules-based medical knowledge*

*9 Duke Margolis Center for Health Policy, 2019, p. 12.*

*10 Several regulatory and standards efforts to define the "trustworthiness" of medical AI are underway. This paper discusses the concept of trust/trustworthiness but does not attempt to define these terms or to set specific requirements around them. To avoid possible conflict or confusion with those regulatory and standards efforts, the former term ("trust") is used instead of the latter ("trustworthiness") in this paper.*

*11 These quality system practices include but are not limited to design control, input verification, process and output validation, usability testing, and postmarket surveillance.*

demonstration of understanding and capability, but by validating that the technology produces reliable and predictable outputs.

Reliance on predicted outputs, however, may not work for machine learning AI, which moves beyond simply performing automated tasks and begins to edge into the practice of medicine. Utilizing the current practice of requiring *a priori* approval of any significant change, regulators will find it difficult to approve or clear a device for marketing if the rationale and evidence behind its actions are unclear or if the device's performance and outputs change over time. Stakeholders (and the regulations and standards that support them) will need to find ways to expand beyond validated, predictable outputs and also consider competence if we are going to learn how to trust machine learning AI, as well as how much to trust it.

Learning to trust AI is proving to be a difficult task for society as a whole—popular fiction is replete with tales of machines that become self-aware and robots that rebel while some futurists warn us of AI's dangers. This is not surprising—trust is derived from knowledge, and there is much we know we do not know about AI, as well as much we do not know we do not know.

If medical device regulators, clinicians, and patients are going to reap the benefits of machine learning AI, it is critical that an appropriate level of trust in these systems be established by a collaborative regulatory system. A lack of trust in AI could affect its acceptance; if machine learning AI technology is not used, clinicians and patients cannot benefit from the advances and efficiencies it offers. The need for trust is even greater with continuous learning models, where performance will change as more training data becomes available and the system refines itself. Users will naturally be suspicious of any system that gives differing results over time.

Conversely, there is danger in over-trusting AI—believing whatever the technology tells us, regardless of the performance limitations of the system. The propensity to trust too much is exacerbated by the current amount of hype that is setting unrealistically high expectations of the technology's competence.[12]

Most people generally trust mature and complex technologies without completely understanding how they work or function. We fearlessly ride elevators without understanding the complicated system of brakes, counterweights, and safety cables that ensure the elevator cars do not fall, and we use our ATM cards without worrying that withdrawals are correctly recorded or that the banks' computers are emptying our accounts. We trust these technologies not because we think there are no potential risks, but because we believe that these risks are adequately managed by the hidden controls incorporated into the system.

Such controls are not uniformly in place for machine learning AI, however, so the accuracy, safety, and performance of these systems cannot be assumed or taken as a matter of faith. While potentially capable of out-performing humans in terms of deriving correlations and patterns that we cannot empirically detect, machine learning systems do not currently demonstrate a similar ability to understand the contextual meaning of data. In linguistic terms, AI, being driven by formal programs and algorithms, is more adept at syntactic (logic and computational) learning than at semantic (meaning-based) learning.[13] Furthermore, the data sets used in AI learning systems are constrained—restricted either in terms of data sources or in terms of the types of data being processed.

The practical implication of these limitations is that data-driven AI systems are not always able to sufficiently evaluate their own base assumptions or to verify the quality of incoming data. They are, to some degree, fragile—they perform extraordinarily well when their base assumptions are solid and the data used is both accurate and relevant. If, however, there are even small errors or changes in this self-contained

*12 "Gartner Says AI Technologies Will Be in Almost Every New Software Product by 2020" https://www.gartner.com/en/newsroom/press-releases/2017-07-18-gartner-says-ai-technologies-will-be-in-almost-every-new-software-product-by-2020*

*13 For example, idioms and euphemisms are not meant to be taken literally and this presents challenges to Natural Language Processing (NLP) systems. For example, discussions about AI ethics may be "a hot potato" to readers of this paper, but that description would be confusing to NLP software. Humor and sarcasm are also artifacts of our everyday discussions but would be misunderstood by software.*

universe of assumptions and data, then the same systems can fail. AI systems are poor at handling the unknown-unknowns—they do not know what they do not know. Thus, any system that can learn can also mislearn—it can "acquire incorrect knowledge"[14] in a variety of ways.

**Ethical AI:** A number of organizations, corporations, and government bodies have published papers and guidelines on ethical AI. Some of the issues associated with ethical AI include bias, lack of explicability, data privacy, poor accountability (who bears the responsibility for a misdiagnosis), and workforce displacement. Discussion about ethical and responsible AI "is primarily driven by recent advancements in AI technologies, growing adoption, and increasing criticality of AI in business decision-making." However, discussion about AI ethics isn't new, rather, it dates back at least to 1942 when introduced by Isaac Asimov. Still, modern AI presents opportunities while also introducing some novel ethical risks due to large datasets, continuous learning processes, etc.[15]

The European Commission's Ethics Guidelines for Trustworthy AI notes that trustworthy AI should be (1) lawful, (2) ethical, and (3) robust. Ethical AI will respect human autonomy and ethical principles and values, such as prevention of harm, fairness, and explicability.

While these efforts to define and prescribe ethical requirements for AI are critical, in the medical AI domain, devices will still be required to adhere to regulatory requirements around privacy and data confidentiality. Moreover, there is one aspect of where medical AI ethics differs from other AI; for medical AI, the ruling ethos of medicine—the Hippocratic dictate to "first, do no harm"—remains the governing rule that overrides all other ethical considerations.

## 4. DATA MANAGEMENT: DATA QUALITY AND BIAS

Data quality is a key factor in the success or failure of a machine learning system; in fact, data quality is as or more important than the machine learning algorithm. There are two main elements that impact data quality: the dataset and the model. The dataset is sent to the model to learn. It is not feasible for a machine to learn outside of this given dataset, and the size and variability define how easily a model can learn from it. Data scientists therefore play an important role with regard to scaling the algorithm.

AI may fail (became untrustworthy) either because data was not representative or not fit for the task to which it was applied. Therefore, the key to making medical AI more trustworthy is ensuring necessary data quality and confirming that algorithms are sufficiently robust and fit for purpose. In short, ensuring the safety and effectiveness of AI depends on verification of data quality and validation of its suitability for the algorithm model. Furthermore, given that AI has the ability to change over time, the processes of verification and validation cannot be a one-time premarket activity, but instead must continue over the life cycle of a system, from the initial design and clinical substantiation, across its post market use, until decommissioning. Continual assurance of the AI-based device's safety and performance across its life cycle will help regulators, clinicians, and patients gain trust in machine learning AI.

There are many aspects that contribute to data quality, including the completeness, correctness, and appropriateness of the data; annotation; bias; and consistency in labelling of the data (e.g., different labels may mean the same thing but the algorithm treats them differently).

*14 Adapted from the Merriam Webster definition of "Mislearn"*

*15 https://www2.deloitte.com/insights/us/en/focus/signals-for-strategists/ethical-artificial-intelligence.html*

Dataset annotations involve variables and biases that humans apply so that an AI solution can spot it.

Any bias that exists in a dataset will affect the performance of the machine learning system. There are many sources, including population, frequency, and instrumentation bias.

Having a system that is unintentionally biased towards one subset of a patient population can result in poor model performance when faced with a different subset, and ultimately this can lead to healthcare inequities. When working with quality data, instances of intentional bias (also known as positive bias) can be present, such as a dataset made up of people only over the age of 70 to look at age-related health concerns.

When considering the application of a dataset for a machine learning application, it is important to understand the claims that it makes. This can be in terms of whether a proper balance in the representative population classes has been achieved, along with whether the data can be reproduced, and if any annotations are reliable. For example, a dataset could contain chest X-rays from males aged 18–30 in a specific country, half of whom have pneumonia. This dataset cannot claim to represent pneumonia in females. It may not be able to claim to represent young males of a particular ethnic group, as this subgroup might not be listed within the dataset variables and might not be plausibly represented in the sample size.

The AI model is trained on the dataset. It will learn the variables and annotations that the dataset is trained on. In healthcare, the vast majority of neural networks are initially trained on a dataset, evaluated for accuracy, and then used for inference (e.g., by running the model on new images).

It is important to understand what the model can reliably identify (e.g., the model claims). Neural networks can generalize a bit, allowing them to learn things slightly different from their training dataset. For example, a model that is carefully trained on male chest X-rays may also perform well on the female population, or with different X-rays equipment. The only way to verify this is to present the trained model with a new test dataset. Depending on the model performance, it may be possible to demonstrate that the AI can accurately identify pneumonia across male and female patients and generalize across different X-ray machines. There may be minor differences in performance between datasets, but these could still be more accurate than a human.

In summary, AI will learn the variables, biases, and annotations of a dataset, with the expectation that it can spot an important feature. Once trained, an algorithm will be tested, revealing that it is able to identify this feature with a certain level of accuracy. In order to test the claim that the AI can identify a specific item, it needs to be tested on a dataset that claims to represent this feature fairly. If it performs to a satisfactory level of performance on this dataset, the model can then claim to be able to identify this item in future datasets that share the same variable as the test dataset.

Figure 2 explains this in more detail.

The following examples show instances where poor quality datasets and their incorrect relationships with algorithm models have caused a failure in the outputs.

An adaptive learning classifier system[16] analyzed photographs to differentiate between wolves and huskies. Instead of detecting distinguishing features between the two canine breeds, the system determined the most salient distinction was that photos of huskies included snow in the background, whereas photos of wolves did not. The system's conclusions were correct with respect to its training data but were not usable in real-world scenarios, because extraneous and inappropriate variables

*16 https://arxiv.org/pdf/1602.04938.pdf*

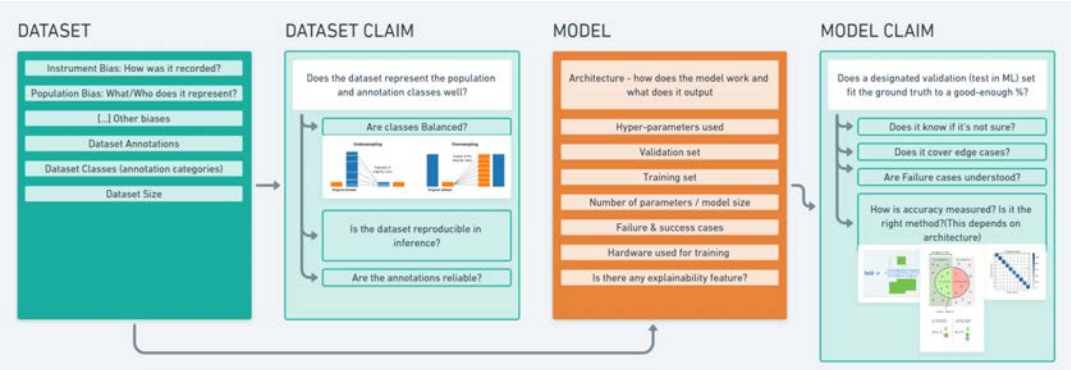


**Figure 2. Data Quality**[17]

(i.e., the backgrounds) were included in the learning dataset.[18] This is an example of how an AI system may detect incidental patterns or correlations in a dataset and assign a causal or meaningful relationship that is incorrect or irrelevant.

Two other recent examples include IBM's Watson using profanity after incorporating the Urban Dictionary into its knowledge-base,[19] and Microsoft's interactive AI assistant "Tay," designed to learn from its interactions with users, which had to be disabled after it was tricked into spouting racist dogma by online pranksters.[20] AI is vulnerable to bad data; it cannot always reliably evaluate the quality of incoming data to determine if it might be biased, incorrect, or invalid. While AI system engineers can create filters to curate data, those filters require assumptions and a *priori* knowledge of the nature and quality of the data. When the assumptions are incorrect and/or the knowledge is insufficient, system performance will be detrimentally affected.

During the devastating California wildfires of 2017, a driving app designed to help users avoid traffic directed fleeing drivers into areas where the inferno was raging as there was less traffic along those routes.[21] Although the system operated correctly for its original purpose of avoiding traffic jams, when that purpose expanded to the more critical function of escaping a wildfire, it did not have adequate information—or sufficiently robust algorithms—to make safe and accurate recommendations. In this instance, the AI system made wrong decisions when it was not correctly matched to the task at hand.

## 5. POTENTIAL REGULATORY AND STANDARDIZATION APPROACHES TO ADDRESS SAFETY AND PERFORMANCE OF AI

As further advancements are made with AI technology, regulators may consider multiple approaches for addressing the safety and effectiveness of AI in healthcare, including how international standards and other best practices are currently used to support the regulation of medical software, along with differences and gaps that will need to be addressed for AI solutions. A key aspect will be the need to generate real-world clinical evidence for AI throughout its life cycle, and the potential for additional clinical evidence to support adaptive systems.

In the last ten years, regulatory guidance and international standards have emerged for software, either as a standalone medical device or where it is incorporated into a physical device. This has provided requirements and guidance for software manufacturers to demonstrate compliance to medical device regulations and to place their products on the market.

However, AI potentially introduces new risk, as discussed in clause 15 of this paper, that is not currently addressed within the current portfolio of standards and guidance for software. Different approaches will be required to ensure the safety

*17 Figure used with permission. Alberto Rizzoli, V7.*

*18 A similar medical AI example occurred when Stanford researchers tested an AI tool to identify melanomas from pictures of moles and found the tool used the presence of rulers in the photos as a positive indicator of cancer. See http://stanmed.stanford.edu/2018summer/artificial-intelligence-puts-humanity-health-care.html*

*19 https://www.theatlantic.com/technology/archive/2013/01/ibms-watson-memorized-the-entire-urban-dictionary-then-his-overlords-had-to-delete-it/267047/*

*20 https://www.nytimes.com/2016/03/25/technology/microsoft-created-a-twitter-bot-to-learn-from-users-it-quickly-became-a-racist-jerk.html*

*21 http://www.latimes.com/local/california/la-me-southern-california-wildfires-live-firefighters-attempt-to-contain-bel-air-1512605377-htmlstory.html*

and performance of AI solutions placed on the market. As these new approaches are being defined, the current regulatory landscape for software should be considered as a good starting point.

In Europe, the Medical Device Regulation (MDR) and In Vitro Diagnostic Regulation (IVDR) include several generic requirements that can apply to software. These consist of the following:

- general obligations of manufacturers, such as risk management, clinical performance evaluation, quality management, technical documentation, unique device identification, postmarket surveillance and corrective actions;
- requirements regarding design and manufacture, including construction of devices, interaction with the environment, diagnostic and measuring functions, active and connected devices; and
- information supplied with the device, such as labelling and instructions for use.

In addition, the EU regulations contain requirements that are specific to software. These include avoidance of negative interactions between software and the IT environment, and requirements for electronic programmable systems.

In the U.S., the FDA recently published a discussion paper[22] for a proposed regulatory framework for modifications to AI/machine learning-based SaMD. It is based upon practices from current FDA premarket programs, including 510(k), De Novo, and Premarket Approval (PMA) pathways. It utilizes risk categorization principles from the IMDRF, along with the FDA benefit-risk framework, risk management principles in the software modifications guidance, and the Total Product Life Cycle (TPLC) approach from the FDA Digital Health Pre-Cert program.

Elsewhere, other countries are beginning to develop and publish papers relating to regulatory guidance. In China, the National Medical Products Administration (NMPA)[23] has produced a guideline for aided decision-making medical device software using deep learning techniques. Japanese and South Korean regulatory bodies have also published guidance for AI in healthcare.

## 6. INNOVATION IN REGULATORY APPROACHES

Modifications to regulatory approaches for AI–based medical device software will depend on the type and nature of the algorithm, and the associated risks. There are existing principles for categorizing SaMD that should form a basis for considering these different approaches.

IMDRF software classification is dependent upon the state of the healthcare condition (critical, serious, or non-serious) and the significance of the information provided by the software (to treat or diagnose, drive clinical management, or inform clinical management). In addition, the international standard IEC 62304[24] introduces three classes of software (A, B, and C), based upon whether a hazardous situation could arise from failure of the software and the severity of injury that is possible.

The level of adaptation of an AI solution also will be important for considering the regulatory approach. As discussed in Clause 2, rules-based AI systems can generally be treated in the same way as traditional software, whereas locked or continuously learning data-driven AI systems will need innovative treatment. The FDA discussion document mentions all currently approved AI solutions have been locked while providing patient care, but there is an ambition to utilize continuous learning systems within the healthcare sector in the future.

A product life cycle approach to regulating AI will be able to allow rapid improvement cycles to software while providing appropriate safeguards. This section

*22 Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning [AI/ML]-Based Software as a Medical Device [SaMD]*

*23 https://chinameddevice.com/china-cfda-ai-software-guideline/*

*24 IEC 62304:2006, Medical device software – Software life cycle processes. 2006.*

will consider the design, development, maintenance, updating, and postmarket activities for AI solutions throughout their life cycle.

Collaboration and coproduction between developers, healthcare providers, academia, patients, governments, and statutory bodies across the AI life cycle will be essential for maximizing the deployment of AI. A recent article from *Harvard Business Review* (July 2019) discussed a concept of "AI marketplaces" for radiology. These are aimed at allowing discovery, distribution, and monetization of AI models, as well as providing feedback between users and developers. Similar collaborations could support the life cycle requirements for AI models, and therefore we recommend establishment of a relationship with IMDRF to develop standardized terminologies, guidance, and good regulatory practices.

» **Recommendation 1:** Working with IMDRF and other regulatory bodies, AAMI and BSI propose development of standardized critical terminology and a taxonomy for medical AI that can inform future national and regional regulatory approaches to the technology.

FDA is currently collaborating with stakeholders to build a U.S. National Evaluation System for health Technology (NEST).[25] This is aimed at generating better evidence for medical devices in a more efficient manner. It will utilize real-world evidence and advanced analytics of data that is gathered from different sources.

Similarly, in the UK, new evidence standards have been developed to ensure digital health technologies are clinically effective and offer economic value.[26] This improves the understanding for innovators and commissioners about what good levels of evidence should look like.

The impact of AI beyond the traditional boundaries of medical device regulation will also be an important factor; particularly where AI is applied in research, health administration, and general wellness scenarios. Alignment with other regulators, e.g., for professional practice, clinical services, research, and privacy will be critical to ensure successful deployment across the healthcare system. The IMDRF is well-suited as the venue to host such discussions and develop related potential regulatory approaches.

The International Medical Device Regulators Forum (IMDRF) is a voluntary group of medical device regulators from nations and regions around the world who have come together to accelerate international medical device regulatory harmonization and convergence by publishing position papers and regulatory guidance and good practices. The IMDRF has also published a document, "Essential Principles of Safety and Performance of Medical Devices and IVD Medical Devices," which details basic requirements for medical device safety and effectiveness.

Due to the potential for AI solutions to learn and adapt in real time, organizational-based approaches to establish the capabilities of software developers to respond to real-world AI performance could become crucial. These approaches are already being considered by U.S. FDA, although they may not necessarily align with EU Medical Device Regulation.

## 7. DEVELOPMENTS IN AI

Good AI development processes and practices will be critical for ensuring the safety and performance of AI solutions in healthcare. These practices will need to address product

*25 https://www.fda.gov/about-fda/cdrh-reports/national-evaluation-system-health-technology-nest*

*26 https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/digital-evidence-standards-framework.pdf*

robustness, algorithm training, validation and testing, modification procedures, and identification and documentation of different versions of an AI solution.[27]

Overall system requirements for safety and security of health software are set out in IEC 82304-1.[28] These requirements are aimed at software products placed on general computing platforms without dedicated hardware, and cover the entire life cycle from design, development, validation, installation, maintenance to disposal of products.

In addition, IEC 62304 covers the software life cycle for medical device software. It applies to software that is regulated within the scope of medical device regulations, and can apply to either standalone software, or software that is embedded into a physical device.

IEC 62304 provides requirements for 'software of unknown provenance' (SOUP): generally available software that has either not been developed for use within a medical device, or for which adequate records of the development process are not available. The additional controls for addressing SOUP in IEC 62304 may provide a starting point for addressing the black-box nature of AI. However, consideration should be given to whether these are for purpose, or if further guidance is required.

There may be some useful AI development practices available from other sources. For example, ISO/IEC JTC1/SC42 is an international committee for generic AI standardization and is currently developing best practices for risk management, bias, trustworthiness, and governance implications. Further standards, best practices, and guidelines are under development in IEEE[29] and ITU/WHO.[30]

An important development aspect will be the definition of the type of AI being used, and its attributes and characteristics that are relevant to regulation and governance. AI is a broadly used term that describes a number of different software technologies. In order to ensure transparency and to drive the correct approaches for safety and effectiveness, it will be important to build a clear understanding of how AI can function. See Clause 11.

## 8. SOFTWARE CHANGES (CHANGE MANAGEMENT)

Requirements for establishing a software change (modification or maintenance) process, including planning, analysis, and implementation are set out in IEC 62304. The standard also provides a process for configuration management (unique identification, change control, history), which, for reasons described previously, may in some instances need to be adapted to meet the unique needs of AI systems. The developer will need to adapt the requirements to suit the needs of the AI solution.

The FDA discussion paper on a proposed regulatory framework for modifications to AI/machine learning-based SaMD suggested modifications to AI would most likely fall under the following categories:

- changes to clinical or analytical performance of the AI, such as increased sensitivity of detecting a condition;
- responses to new data inputs (e.g., compatibility with other sources of the same input data type, or expansion of the types of input data utilized within an AI solution); and
- alterations related to the intended use of the software that are claimed by the developer, that result in a change to the significance of the information provided by the AI, or a change in the healthcare situation.

The regulatory approach for modifications to AI software will be dependent on the extent of these changes, and potentially by the way in which modifications are

*27 See references to Good Machine Learning Practices (GMLP) in the FDA discussion paper (Footnote 1)*
*28 IEC 82304-1:2016, Health software – Part 1: General requirements for product safety. 2016.*
*29 https://ieeexplore.ieee.org/document/8442729*
*30 https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/default.aspx*

anticipated prior to the changes being made. It may be possible for an AI developer to specify any modifications that they plan to achieve in the future (i.e., once the AI is in use). The developer would need methods in place to achieve any anticipated changes control any risks associated with them.

There will be limitations to the scope of anticipated changes that can be specified in advance, depending upon how extensive the modifications will be. However, the ability to monitor AI performance in real time provides an opportunity to develop a dynamic regulatory process that allows rapid improvements to the software while ensuring safety.

## 9. FURTHER QUALITY AND RISK MANAGEMENT CONSIDERATIONS

The differences between AI and traditional software have been identified in earlier AAMI-BSI workshops and are summarized in this whitepaper. The impact of these differences on quality and risk management processes and on systems is summarized below.

The input datasets required to test and train algorithms will need to be predefined, relevant, and appropriate. Data will need to be provided in sufficient volume, variety, and accuracy to ensure that the algorithm can learn effectively. Adequate checks must be madeto be made to ensure that the representation of input data will be satisfactory for ensuring the overall safety and performance of the AI.

Validation of an AI solution to ensure that it meets its intended use and the needs of the user will be more complex when compared to traditional SaMD. Likewise, the verification process for adaptive algorithms will not be the same when compared to rules-based software. This is because AI has the ability to respond differently to particular data inputs over time, and so the outputs cannot easily be predicted. Proof-of-concept studies are underway to generate and evaluate synthetic healthcare data for the purposes of validating machine learning algorithms.[31] This could provide a number of benefits, including mitigating bias, providing ability to benchmark different AI solutions against a common dataset, and reducing costs and privacy issues relating to data generation.

Performance metrics of algorithms will be an important factor for developers to consider. This will allow real-time monitoring of AI solutions against their predicted outputs. The ability to quickly identify and react to real-time outcomes is an essential element for a SaMD solution. Adapting to real-world performance metrics allows developers to continuously monitor and improve on marketed AI solutions and is important in gaining public trust.

Explicability of AI outputs, including the level of supervision that an AI solution utilizes in its learning process will be an essential aspect in ensuring safety and performance.

Supervised learning involves how input variables (data) map to a particular set of outputs. However, unsupervised learning is used to infer patterns from data without reference to known or labelled outcomes.

"Supervised Learning" is a common but often misunderstood term. When used in a machine learning context, it means that the software maps an input to an output based on labelled data training. It does not mean that there is a human supervisor overseeing the software. "Unsupervised learning" uses a model to learn patterns from un-labelled data, without any predicted output variables.

*31 https://ieeexplore.ieee.org/abstract/document/8787436*

The level of autonomy provided by an AI solution will need to be considered, both from a human factor approach and also any potential impact on liability and/or professional practice regulations. Another important factor will be the degree of clinical oversight that is provided to a continuous learning system that is providing patient care in real time.

Reducing the risk of bias within AI solutions will be a further consideration for developers. Bias can be introduced through machine-related aspects, such as incorrect application of datasets or the wrong algorithm model. There is also the possibility of bias being introduced though human or institutional interventions, such as using training data from narrowly selected demographics, clinical cases, or treatment protocols. Bias can be amplified by AI processes. However, a properly designed system can minimize or reduce bias over time, through introduction of new and varied data sets.

## 10. POSTMARKET SURVEILLANCE

Medical device manufacturers are already obliged to undertake postmarket surveillance activities for regulatory purposes. However, the resolution process for problems relating to AI is likely to add complexity, due to their adaptive nature and also the lack of proper understanding of their internal workings. This makes the methodologies for undertaking root cause analysis on an AI solution difficult. Transparency around the function of an AI solution, along with any modifications undertaken will be a key safety aspect that will also help to drive adoption.

As previously discussed, SaMD solutions are in a unique position to build in mechanisms that quickly identify safety or effectiveness concerns through real-world performance monitoring. This real-world performance monitoring is a key strategy for AI solution postmarket surveillance.

## 11. TERMINOLOGY AND TAXONOMIES

Defining AI has proven to be a complicated endeavor. This paper concentrates on how a specific type or aspect of AI—machine learning—can be addressed by standards and regulations, but there are many different and divergent types and definitions of AI. Initial efforts to define terms for AI are underway[32]; Annex A of this whitepaper lists selected definitions for critical terms used in this document.

Differing definitions or taxonomies of AI by various regulatory authorities will create inefficiencies and confusion for medical device manufacturers and could hinder the development and adoption of medical AI. This will also impede the development of standards to support that regulation and promote medical AI safety and efficacy.

It is essential that national or regulatory authorities adopt consistent terminologies and taxonomies for AI in medical technologies. Stakeholders in cooperation with regulators, such as the U.S. FDA, UK MHRA, and the IMDRF, must identify and define critical terminology and develop a taxonomy of AI that can inform national and regional authorities as they develop their own approaches to regulatory medical AI. (See Recommendation 1).

A further recommendation is that IMDRF establish a working group to address issues around AI in healthcare to prepare needed guidance and good regulatory practices in AI.

» **Recommendation 2:** AAMI and BSI recommend that IMDRF establish an AI working group to address issues around AI in healthcare and to prepare needed guidance and good regulatory practices.

*32 The Consumer Technology Association has published a standard defining terms related to AI and associated healthcare technologies (ANSI/CTA 2089.1, Definitions/Characteristics of Artificial Intelligence in Health Care). The ISO/IEC JTC1/SC42 Artificial Intelligence committee is developing ISO/IEC 22989 "Information Technology – Artificial Intelligence – Artificial Intelligence Concepts and Terminology."*

## 12. LIFE CYCLE MAPPING AND GUIDANCE

Medical device regulatory standards often address horizontal principles that apply to many types of products/software (e.g., usability) or process requirements throughout the life cycle (e.g., quality, risk management). Where international regulatory standards already exist, guidance should be developed relating to their application for AI solutions. Additionally, gaps in the international standards landscape should be identified, so appropriate guidelines can be developed.

» **Recommendation 3:** AAMI and BSI recommend mapping AI-applicable international regulatory standards (where such exist) to the October 2018 IMDRF Essential Principles and identifying gaps where additional new standards or guidance are needed.

An overarching "umbrella" standard that sets out references to all of the requirements/recommendations for the safe and effective deployment of AI within a healthcare system could provide a useful overview. Such a standard would describe the key principles that need to be addressed across the AI life cycle, from the perspectives of developers and the healthcare system. This overarching guidance would reference existing standards and best practices rather than create a new set of requirements, but it would become a single document that provides a clear set of instructions for what to consider. It could also be used to map IMDRF essential principles against existing standards.

## 13. GUIDANCE AROUND DATA QUALITY AND MAINTENANCE

There is a need for additional information regarding factors that affect data quality in regard to AI. However, an initial scoping exercise and research should take place to ensure that any guidance remains relevant to the regulation of AI as a medical device, while addressing any relevant needs across the supply chain.

» **Recommendation 4:** AAMI and BSI recommend developing guidance on factors affecting data quality in regard to AI as a medical technology.

BSI and AAMI acknowledge there are many factors that can have an impact on the quality of data used in AI and a significant number of initiatives are working to address these challenges. Some of these have been highlighted earlier within this whitepaper, including dataset size, annotations, and biases. Other factors could also be applicable for data quality that is applied in regulated situations. For example, data storage could be an attractive target for hackers or, if a storage solution allows data to be corrupted, then the performance of AI that depends on that data would be adversely affected.

## 14. AI EVALUATION PROTOCOL

Deployment of AI in healthcare is currently being explored from many other perspectives beyond regulatory approval. These include organizational management, professional conduct, research and ethics, evidence-based practice, and data governance.

A comparison of the best practice recommendations within each of these perspectives reveals a degree of overlap, and by identifying these commonalities it should be possible to develop a common set of criteria or questions that could be used as an evaluation protocol by multiple stakeholders. This could include the following:

- Identifying patient benefit from the technology
- Patient safety and security
- Data curation and accessibility
- Clinical association and validation
- Performance metrics and health outcomes
- Transparency, equality, bias, and acceptability
- Routine monitoring of continuous learning systems
- Cost effectiveness and fair commercialization
- Doctor–patient–machine relationships.

» **Recommendation 5:** AAMI and BSI recommend establishing a common criteria for the deployment of AI in healthcare systems that could be used as an evaluation protocol by multiple stakeholders, covering organizational management, professional conduct, research and ethics, evidence-based practice, and data governance.

## 15. RISK MANAGEMENT/BASIC SAFETY GUIDANCE

Whereas the management processes and core activities for risk analysis, risk evaluation, risk control, and evaluation of overall residual risk will remain the same, data-driven AI systems will introduce new failure modes and hazards. These include increased levels of autonomy, reducing the risk controls requiring human intervention, and the "black box" nature of some AI systems making quality assurance difficult.

There is need for guidance on risk management for AI as a medical technology. Such guidance should cover different failure modes and hazards that are unique to AI systems. The guidance should identify specific considerations that AI developers should examine when applying the requirements of ISO 14971 to an algorithm.

» **Recommendation 6:** AAMI and BSI recommend developing risk management guidance to assist in applying ISO 14971 to AI as a medical technology.

## 16. VALIDATION VS. COMPETENCIES (GUIDANCE)

There is a need for information regarding validation of AI systems. Due to the opaqueness of many machine learning systems, there will be an increased reliance on validation studies to demonstrate the performance and accuracy of machine learning solutions. Although there is existing guidance related to the key characteristics of validation study design, execution, and evaluation, the commercial adoption of a machine learning solution may be highly dependent on the performance and limitations of the product.

» **Recommendation 7:** AAMI and BSI recommend developing guidance on factors to consider in the validation of AI systems and on the use of nontraditional approaches, such as excellence assessments, to demonstrate a reasonable assurance of product quality.

This guidance will discuss various performance characteristics (e.g., sensitivity, specificity), presentation approaches (e.g., ROC curves, Confusion Matrix), and the role of benefit-risk evaluations as means to communicate product performance. As adaptive systems may require multiple validations, such guidance would also discuss methods to streamline the execution of validation studies.

 

Even with enhanced validation study guidance, there will always be a risk of bias in a validation study that may result in artificially high performance. One way to mitigate this is by using good, and possibly excellent, development processes. A thoughtful and pragmatic development process is more likely to create good software than a compliance-only based development process. Product quality is related to the process used to develop the product, and this has been noted in the U.S. FDA's proposed Pre-Certification program.[33]

**Validity versus validation:** The term validation has a special meaning in both the medical device world and the data science world. For medical devices, validation is a process that is used to ensure that user needs are met. For data science, validation is a process to ensure that the data has validity (i.e., that the data is correct and adequate for its intended purpose).

*33 https://www.fda.gov/medical-devices/digital-health/digital-health-software-precertification-pre-cert-program*

# ANNEX A – GLOSSARY

This Annex does not intend to provide a comprehensive list of all terms associated with the topic of healthcare artificial intelligence. Rather, the terms defined in this Annex are intended to be informative towards this whitepaper.

**A.1**

**algorithm**

a process or set of rules, including data driven or human-curated, to be followed in calculation or other problem-solving operations. The technology of artificial intelligence uses a variety of algorithms as tools and applications.

[Source: ANSI/CTA-2089.1]

**A.2**

**Artificial Intelligence (AI)**

**(1)** <system>capability of an engineered system to acquire, process, and apply knowledge and skills.

Note 1: Knowledge are facts, information, and skills acquired through experience or education.

[Source: SC42, draft 22989]

**(2)** A machine's ability to make decisions and perform tasks that simulate human intelligence and behavior.

[Source: *Xavier Health, Perspectives and Good Practices for AI and Continuously Learning Systems in Healthcare*]

**(3)** A general term addressing machine behavior and function that exhibits the intelligence and behavior of humans.

[Source: ANSI/CTA-2089.1]

**AI Beyond Artificial: Assisted, Augmented, and Autonomous:** Three other terms often come up when discussing artificial intelligence: assisted, augmented, and autonomous intelligence. PricewaterhouseCoopers broadly separates these three terms as helping people perform tasks faster (assisted intelligence); helping people make better decisions (augmented intelligence); and automating decision-making processes without human interventions (autonomous intelligence) .

The term "augmented intelligence" is sometimes used instead of Artificial Intelligence to emphasize how the technology enhances rather than replaces human intelligence.

**A.2**

**bias**

favoritism towards some things, people, or groups over others.

[Source: ISO 24027]

**A.3**

**continuous learning**

incremental training of an AI system that takes place on an ongoing basis while the system is running in production.

[Source: SC42, draft 22989]

**A.4**

**deep learning**

**(1)** approach to creating rich hierarchical representations through the training of neural networks with many hidden layers.

Note 1: Deep learning uses multilayered networks of simple computing units (or "neurons"). In these neural networks each unit combines a set of input values to produce an output value, which in turn is passed on to other neurons downstream.

[Source: SC42, draft 22989 references ISO/IEC 23053, 3.13]

**(2)** an advanced form of neural network machine learning that utilizes big data to generate impressive results.

[Source: CTA, What is Artificial Intelligence?]

**A.5**

**effectiveness**

reasonable assurance that a device is effective when it can be determined, based upon valid scientific evidence, that in a significant portion of the target population, the use of the device for its intended uses and conditions of use, when accompanied by adequate directions for use and warnings against unsafe use, will provide clinically significant results.

["21 Code of Federal 860.7)].

**A.5**

**machine learning**

**(1)** function of a system that can learn from input data instead of strictly following a set of specific instructions.

Note 1: MACHINE LEARNING focuses on prediction based on known properties learned from the input data.

[Source: AAMI TIR66, Guidance for the creation of physiologic data and waveform databases to demonstrate reasonable assurance of the safety and effectiveness of alarm system algorithms]

**(2)** a sub-branch of AI in which the rules by which a decision or action are taken are learned through examples, a training process.

[Source adapted: BSI, Recent advancements in AI – implications for medical device technology and certification]

**A.6**

**Natural Language Processing (NLP)**

**(1)** information processing based upon natural-language understanding.

Note 1: NLP is a field of AI.

Note 2: Natural language is any human language, such as English, Spanish, Arabic, or Japanese, to be distinguished from formal languages, such as Java, Fortran, C++, or First-Order Logic.

Note 3: Examples of natural language are text, speech, gestures and sign language.

[Source: SC42, draft 22989]

**(2)** an application of AI, computer science, and information engineering by which the technology can understand written or spoken human conversation.

[Source: ANSI/CTA-2089.1]

### A.7

### neural network/neural net/artificial neural network

network of primitive processing elements connected by weighted links with adjustable weights, in which each element produces a value by applying a nonlinear function to its input values, and transmits it to other elements or presents it as an output value.

Note 1: Whereas some neural networks are intended to simulate the functioning of neurons in the nervous system, most neural networks are used in artificial intelligence as realizations of the connectionist model.

Note 2: Examples of nonlinear functions are a threshold function, a sigmoid function, and a polynomial function.

[Source: SC42, draft 22989, references ISO/IEC 2382-28:1995]

### A.8

### postmarket surveillance

systematic process to collect and analyze experience gained from medical devices that have been placed on the market.

[Source ISO 13485:2016, 3.14]

### A.9

### robustness

ability of a system to maintain its level of performance under any circumstances.

[Source: SC 42, draft 22989]

### A.10

### Software as a Medical Device (SaMD)

software intended to be used for one or more medical purposes and to perform these purposes without being integral to the hardware of a medical device.

[Source: IMDRF, Software as a Medical Device (SaMD): Key Definitions]

### A.11

### training

process to establish or to improve the parameters of a machine learning model, based on a machine learning algorithm, by using training data.

Note 1: For supervised learning, the machine learning model can be trained (learn from) data that is similar to input data.

**Note 2:** For transfer learning, the input data is not necessarily similar to the training data.

**Note 3:** For unsupervised learning, the machine learning model is trained (learns from) and makes inferences, or predictions, based on the same data.

[Source SC42, draft 22989 references ISO/IEC 23053, 3.9]

### A.12

**transparency**

open, comprehensive, accessible, clear, and understandable presentation of information.

[Source: ISO 20294:2018, 3.3.11]

### A.13

**validation**

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.

[Source: ISO 9000:2015, 3.8.13]

### A.14

**verification**

**(1)** confirmation, through the provision of objective evidence, that the specified requirements have been fulfilled.

**Note 1:** Verification only provides assurance that a product conforms to its specification.

[Source: ISO/IEC 27042:2015, 3.21]

**(2)** confirmation by examination and, through the provision of objective evidence that specified requirements have been fulfilled.

**Note 1:** The term verified is used to designate the corresponding status.

**Note 2:** Confirmation can comprise activities such as:

- performing alternative calculations;
- comparing a new design specification with a similar proven design specification;
- undertaking tests and demonstrations;
- reviewing documents prior to issue.

[Source: IEC 60601-1:2005+AMD1:2001 [42], definition 3.138]


医课汇
公众号
专业医疗器械资讯平台
WECHAT OF
HLONGMED
hlongmed.com
医疗器械咨询服务
MEDICAL DEVICE
CONSULTING
SERVICES
医课培训平台
医疗器械任职培训
WEB TRAINING
CENTER
医械宝
医疗器械知识平台
KNOWLEDG
ECENTEROF
MEDICAL DEVICE
MDCPP.COM
医械云专业平台
KNOWLEDG
ECENTEROF MEDICAL
DEVICE
AAMI
dvancing Safety in Health Technology
www.aami.org