
# Cybersecurity Risk Analysis for Medical Devices in the Era of Evolving Technologies

April 2026



MITRE

# Notice

The views, opinions and/or findings contained in this report are those of The MITRE Corporation and should not be construed as an official government position, policy, or decision, unless designated by other documentation.

This technical data report was produced for the U.S. Government under Contract Number 75FCMC23D0004, and is subject to Federal Acquisition Regulation Clause 52.227-14, Rights in Data-General.

No other use other than that granted to the U.S. Government, or to those acting on behalf of the U.S. Government under that Clause is authorized without the express written permission of The MITRE Corporation.

The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services.

For further information, please contact The MITRE Corporation, Contracts Management Office, 7515 Colshire Drive, McLean, VA  22102-7539, (703) 983-6000.



# Authors

Melissa P. Chase
Steven Christey Coley
Moses Liskov
Margie Zuk

The MITRE Corporation, Bedford, MA

# Acknowledgments

MITRE would like to thank the medical device manufacturers, healthcare delivery organizations, cybersecurity and regulatory consultants, and cybersecurity vendors that provided insights into managing cybersecurity risks when designing medical devices to incorporate emerging and evolving technologies. The considerations and resources provided in this document are a direct result of lessons learned from these engagements.

# Table of Contents



# Table of Figures

# 1 Introduction

As Medical Device Manufacturers (MDMs) continue to innovate, the technologies used to provide new capabilities and to protect devices against existing and emerging threats evolve. MDMs may start to adopt relatively established technologies or to incorporate emerging technologies, which present new cybersecurity risks that may impact device functionality and lead to patient harm if not addressed.

This discussion paper addresses some general cybersecurity considerations when designing medical devices to incorporate emerging and evolving technologies. It then considers three specific examples: cloud computing, artificial intelligence and machine learning (AI/ML), and post-quantum cryptography (PQC). Cloud computing and AI/ML are evolving technologies that present new cybersecurity risks, while PQC is an emerging technology to manage the risks posed by quantum cryptography.

MITRE developed this paper by conducting a landscape analysis that included a literature review and interviews with a broad sample of stakeholders, including MDMs, Healthcare Delivery Organizations (HDOs), cybersecurity and regulatory consultants, and cybersecurity vendors. MITRE identified the key cybersecurity challenges for these technologies facing MDMs and mitigations and consensus practices.

# 2 General Considerations

MDMs bring healthcare innovations to patient care in part by leveraging emerging and evolving technologies. When developing approaches to managing the cybersecurity risks introduced by these technologies, it is useful to consider the wide range of medical device characteristics and the variety of contexts in which they are used.

Medical devices range from implantable devices (e.g., pacemakers) to small devices intended to be portable (e.g., infusion devices), or to devices that are large systems integrated into a hospital's clinical environment (e.g., MRIs and CT scanners). Many of these devices have limits on power, memory, and computational power, which may constrain the cybersecurity algorithms that can be used as cybersecurity controls. Large devices typically have long lifetimes and may run outdated hardware and software. In addition, most devices interface with other medical devices, products, and health IT systems, which introduce additional cybersecurity risk.

The environments in which devices are used are evolving from being primarily used in healthcare facilities (e.g., hospitals, nursing homes) to ambulatory environments in which devices are managed by the patient's healthcare providers (e.g., home dialysis) to devices usually used and managed by the patients such as home blood pressure monitors and wearable devices with connectivity to apps. As devices move outside healthcare facilities, HDO staff have less control over these devices and how they are managed.

To accommodate those changes, risk management paradigms for evolving technologies, such as cloud services and AI/ML, will account for medical devices and components that may be located outside of a healthcare facility and managed by third parties. Cybersecurity risk management of medical devices has always been a shared responsibility between MDMs and HDOs, but now additional third parties are accounted for when defining roles and responsibilities.

A key consideration is designing devices and putting process and frameworks in place to prevent devices from becoming devices “that cannot be reasonably protected against current cybersecurity threats” [1], such that they are not cybersecure. Medical devices and their systems can evolve through hardware and software upgrades over the lifetime of the device, so they can continue to be used safely. Threat modeling early in design and development can lead to robust, evolvable designs and threat models with appropriate mitigations. These practices, in turn, can help ensure that vulnerabilities discovered in the post-market are more likely to be controlled, or if uncontrolled, facilitate rapid remediation.

# 3 Cloud

## 3.1 Overview

Cloud computing has its roots in 1960s time-shared systems, and modern cloud “Infrastructure as a Service (IaaS)” systems were first stood up in the early 2000s. Over the years, medical devices have incorporated cloud computing in their systems; however, we consider cloud computing as an emerging technology in the context of medical devices, as its use in devices is evolving and cloud technologies are becoming an integral part of a medical device’s essential performance. The integration of cloud-based technologies in medical devices and throughout their lifecycles has significantly increased in recent years, and even more so with the increased adoption of AI/ML technologies that often rely on cloud computing infrastructure.

The National Institute of Standards and Technology (NIST) developed a widely used framework for defining cloud computing [2] that is illustrated in Figure 1.

NIST characterizes cloud computing as “a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction” [2].

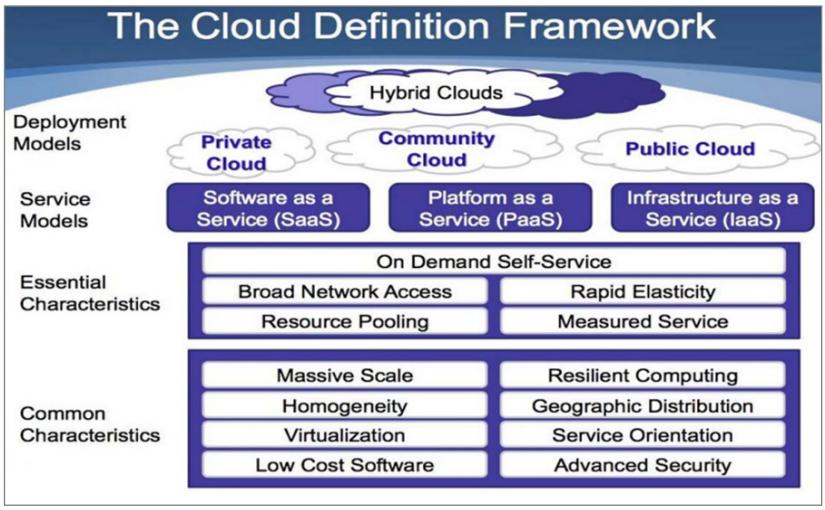


Figure 1: NIST Cloud Definition Framework
Source: Udoh, Patterson, and Cordle [3]

NIST defines three service models suitable for different use cases:

- **Software as a Service (SaaS).** The customer uses the cloud provider's applications on the cloud infrastructure, with limited to no control over the infrastructure, Operating System (OS), or storage.
- **Platform as a Service (PaaS).** The customer runs their applications on the cloud infrastructure, with control over the deployed applications, but limited or no control over the infrastructure, OS, or storage.
- **Infrastructure as a Service (IaaS).** The customer runs applications, OSs, and other software on servers, networks, and other computing resources provided by the cloud provider. The customer has control over the applications, OSs, storage, etc., but not the lower-level computing resources or infrastructure.

Finally, NIST describes multiple deployment models with different trade-offs between control of resources, costs, and scale:

- **Private cloud.** The cloud infrastructure is provisioned to a single organization, managed by that organization or a third-party for the organization, and may exist on or off premises.
- **Community cloud.** The cloud infrastructure is provisioned to a community of organizations with a common interest, managed by one or more of those organizations or a third-party, and may exist on or off premises.
- **Public cloud.** The cloud infrastructure is provisioned for public use, owned and managed by a provider (commercial, non-profit, government, or some combination), and exists on the cloud provider's premises.
- **Hybrid cloud.** A combination of the above deployment models that uses technology that enables data and application portability to support load balancing and other capabilities.

An MDM can use cloud infrastructure in different ways, internally to support device development or as an integral part of the deployed device.

As an example of the first use case to support device development, an MDM might use cloud infrastructure to develop an AI algorithm, which would be part of the medical device; however, the cloud infrastructure would not be included in the fielded device. This can be seen as a one-time development activity. The MDM could use any of the cloud services and any of the deployment models described above to achieve this use case. In this case, the MDM is the customer, and the cloud resources may be provided by the MDM (private cloud), third party providers, or a combination of both.

The second use case shows cloud infrastructure as part of the deployed device. The MDM uses cloud computing to provide resources to its customers, including HDOs, physician practices, or individual patients, for example, collecting data from devices, or delivering information used to guide therapies, or provide diagnoses. The MDM will be a customer of cloud providers and other third parties, and can obtain cloud resources through any of the cloud service models, depending upon how much of the application stack the MDM wants to manage (see Figure 2). From the perspective of the MDM's customers, the MDM is most likely providing these resources as a SaaS system, managing the full application stack.

## 3.2 Challenges

When adopting cloud computing technology, MDMs face challenges in managing their assets on the cloud, including cybersecurity threats and risks to those assets, which in turn may impact patient safety. These challenges fall into two broad categories: acquisition/operational deployment and software development.

### 3.2.1 Cloud Computing Acquisition and Operational Deployment

Providing cloud computing resources to customers (i.e., HDOs, clinical practices, and patients) falls along a spectrum ranging from proprietary software deployed on servers at the customer's site to commercial or open-source software deployed on the MDM's premises or on third-party servers to one of the cloud service types delivered by cloud service providers.

While cloud services may reduce cost and provide easier implementations and efficiency, it comes with the risks related to using third-party services. Once data and services are in the cloud, both MDMs and HDOs have less control over the device and data and have increased their overall cybersecurity risk, including risk to patients. Cloud service providers are not regulated and introduce unknown systemic risks. Downstream impacts of attacks on cloud-based medical device services are not always accounted for, potential choke points aren't always identified, and risk assessments might not be asking the right questions. If cloud services are not available, the medical devices and HDOs may not be able to provide care to patients, which leads to indirect and latent patient harm. As such, developing contingency plans would allow for further preparedness. The purchasing controls specified in clauses 7.4.1–7.4.3 in ISO13485:2016 provide a framework for setting requirements for suppliers, which can include requiring cloud service providers to establish appropriate contingency plans to prevent patient harm. Indeed, an HDO might not be willing to assume the risk of a cloud-based approach and decide to acquire a more traditional system.

MDMs have traditionally delivered their devices to HDOs who assume responsibility for operating these devices on their network; however, due to the complexity of cloud-based systems, the different roles and responsibilities are reassessed. The MDM may become the partial operator of the device, namely the cloud-based component, and that component might be managed by the MDM on-premises or might rely upon a third-party PaaS or a cloud provider's IaaS. In addition, it may be challenging for MDMs to carry out penetration testing, vulnerability scanning, and other cybersecurity testing of third-party cloud resources and services used in the medical device, since they are managed by third parties (i.e., CSPs and container providers) and MDMs will follow rules of engagement for such testing.

Finally, since cloud resources may be provisioned globally, the MDM will likely deal with different data and device regulatory regimes, as well as the complexity of global management, which may constrain solutions for providing encryption, public key infrastructure, multiple locations of data, robust backup strategies, etc.

### 3.2.2 Software Development Approaches

There are different approaches that MDMs may take during the development of their devices. The simplest way is cloud-hosted solutions, where an MDM takes advantage of cloud computing capabilities by containerizing their software, and deploying it to on-premises servers or to a third-party cloud provider. But to take full advantage of the cloud, the MDM has to move from cloud-hosted solutions to cloud-native solutions. Cloud-native solutions rely on solutions for databases and analytics that have been designed from the ground up by cloud providers and third-party developers to run on cloud infrastructure and offer better performance, scaling, and resilience.

Cloud-native solutions call for different development skills, since the application stacks, frameworks, and development environments may not use the same technologies used in developing traditional medical devices. MDMs develop DevSecOps approaches to ensure that development and provisioning are handled securely and develop continuous integration (CI)/continuous deployment (CD) pipelines, which include cloud provisioning actions (e.g., building virtual machine (VM) images, defining clusters) and specific security services (e.g., scanning containers for vulnerabilities, monitoring VMs).

## 3.3 Threats and Risks

The risks that cloud services present to patients vary based on how they are used, as cloud services can be used throughout the lifecycle of a medical device—during design and development, operational deployment of essential functionality, or for providing ancillary services. For example, during development, cloud services may be used to train a machine learning algorithm, which then becomes part of a device or stand-alone software as a medical device (SaMD). An adversary could attack the cloud services used for training (e.g., by tainting the training data) and cause the algorithm to be incorrect, which could lead to misdiagnosis or improper treatment when the algorithm is deployed.

A medical device can be developed as a SaaS, with its essential functionality provided by cloud resources instead of on-premise resources, or the device may use the cloud as a platform to perform ancillary activities, such as collecting data from the device for off-line analysis. If an adversary attacks the cloud resources used by the medical device, the device might be rendered unavailable (e.g., in a ransomware attack that encrypts data used by the device) or might not operate correctly (e.g., data or code is changed), which may lead to misdiagnosis, under or over treatment, or delays in treatment.

Attacks against cloud-based medical devices can have a far-reaching and wide impact and affect multiple HDOs and patients. Cyberattacks against cloud services that provide resources to fielded medical devices can impact devices in different facilities. For example, the ransomware attack against Elekta's cloud services affected cancer treatment at over 170 facilities [4].

Cyberattacks against cloud services during product development can lead to supply chain attacks and disruptions. An attack against cloud infrastructure or CI/CD pipelines may cause the development environment to be unavailable, potentially disrupting the supply chain and slowing the development of new therapeutic and diagnostic procedures. And as described above, an adversary can introduce malware or interfere with the training of the algorithm under development, and these changes would be present in all of the deployed systems.[1]

## 3.4 Mitigations

Mitigations for the challenges and risks described above fall into three categories: policies and processes, resilient architecture and design, and preparedness and response.

**Policies and processes.** Since one of the key challenges is the differing roles and responsibilities between cloud and on-premise deployments, medical devices manufacturers could define these roles based upon the architecture of the cloud-based medical device. Figure 2 illustrates the management responsibilities for the different cloud service models. Although the figure was developed for government agencies, it is informative for managing cybersecurity risk in cloud-based medical devices, both from the perspective of the MDM acting as the vendor and HDO acting as the "agency," and the MDM as the "agency" and the CSP as the vendor. In determining needs, roles, and responsibilities, the MDM considers which service model they are using as provider and as customer, who will be managing each of the layers, and how security will be implemented.

Service Level Agreements (SLAs) and contracting language, between HDOs and MDMs, and MDMs and CSPs, can be used to delineate the responsibilities and define the security expectations. The MDM's enterprise security requirements form the foundation of the cloud service security expectations and will be augmented with their customer's requirements, including defining availability and backup requirements. CalTech's Center for Technology and Management Education's blog has a beginner's guide to Cloud SLAs [5]. NIST's *Cloud Computing Synopsis and Recommendations* [6] discusses SLAs and suggests how to apply NIST SP 800-53 controls "by considering the patterns of different provider and consumer relationships for sharing security responsibilities." Clauses 7.4.1–7.4.3 in ISO13485:2016 provide a framework for purchasing controls that MDMs can use to ensure that the procured cloud services conform to the specified cybersecurity requirements, including evaluating and monitoring suppliers, notification of changes, and maintaining documentation. Some of the clauses in the Healthcare & Public Health Sector Coordinating Councils' model contract language may be adapted to apply to cloud services [7].

[1] Although medical devices aren't mentioned, [20] discusses the supply chain impacts of attacks against CSPs.


Figure 2: Responsibilities for Different Cloud Service Models

*Source: CISA [8]*

Medical devices that are built on cloud services are still composed of systems and software, and the processes for secure development through the product lifecycle apply. But, especially when developing cloud-native SaaS solutions, the DevSecOps approaches will be different, focusing on continuous integration/continuous development pipelines; managing container and orchestration security; and integrating security throughout development, testing, and deployment. MDMs will develop their cloud-based development processes and hire or train staff on them.

**Resilient Architecture and Design.** The processes used to ensure the secure, resilient, and safe design and development of medical devices could also apply to cloud-based medical devices. SBOMs for cloud based medical devices would include all cloud components, such as virtual machines, containers and all the layers in the container image, the machine image, and the cloud-native services. Likewise, threat model development would include the cloud components and services that are part of the device. The *Playbook for Threat Modeling Medical Devices* [9] provides recommendations on applying the four-question threat modeling framework to medical devices. When defining the architecture, including high-value data flows and trust boundaries, consider the cloud technology stack and the different roles and responsibilities for the different service models. When identifying the threats (what can go wrong?) CAVEaT [10] offers some general considerations specific to cloud technologies (e.g., attacks may differ by service model). One of the methodologies for identifying threats is to use cyber lifecycle attack frameworks like ATT&CK. The Enterprise ATT&CK enterprise matrix includes adversary

tactics and techniques against cloud services, including those specific to IaaS and Saas services.[2] MITRE's Center for Threat Informed Defense's Mappings Explorer,[3] which maps AWS, Azure, and Google security controls to ATT&CK tactics and techniques for AWS, Azure, and Google, can help identify appropriate mitigations specific to the different cloud stacks. The Open Web Application Security Project (OWASP) cheat-sheet series includes cheat sheets for secure cloud architecture, secrets management, Docker security, and Kubernetes security.[4] Government agencies, including NIST, CISA, and the National Security Agency (NSA) have also published best practices for cloud security, some of which are listed in Appendix A-2.

Cloud-based medical devices have an increased attack surface because more systems are off-premise and exposed to the Internet. At the same time, cloud services may be easier to secure because CSPs manage the security of lower layers of the cloud infrastructure. In addition, when building machine images and containers, MDMs can use trimmed-down versions of operating systems and platforms to reduce exposure. Finally, CSPs and container providers offer tools to enhance security—including controls for managing secrets—such as passwords and access tokens, observability tools for monitoring the cloud-based systems, and tools for scanning for and managing vulnerabilities.

Preparedness and Response. Although cloud-based medical devices offer the promise of resilience and availability, cloud infrastructure is not immune from cyberattack, which can impact the medical device's operations. Typically, cloud-based systems are provisioned in a region with built-in redundancy to handle server outages. For additional resilience, an MDM could provision their systems across multiple geographic regions. When doing this, it is important to consider regulatory concerns in different regions, for example, involving encryption, to ensure that appropriate security controls are implemented if not provided by the CSP.

In addition to relying upon cloud infrastructure redundancy, MDMs can include in the device architecture the ability to operate when the cloud is unavailable through local caching (perhaps implemented as a hybrid cloud) and creating backups in different locations to facilitate restoring systems that may have been affected by a cyberattack.

[2] https://attack.mitre.org/matrices/enterprise/

[3] https://center-for-threat-informed-defense.github.io/mappings-explorer/

[4] https://cheatsheetseries.owasp.org/

# 4 Artificial Intelligence / Machine Learning (AI/ML)

MDMs and HDOs have a growing interest in integrating AI/ML to improve speed and performance into medical devices, clinical decision support systems, and related products.[5] While there is a widespread belief that AI/ML will have a significant impact on the healthcare industry, adoption of AI/ML has been relatively slow for many types of devices. While there has been an increase in AI-enabled medical devices, based on the interviews that we held, it is not generally clear to MDMs and HDOs where AI/ML would provide benefits that outweigh the risks it may introduce.

## 4.1 Overview

With AI/ML's variety and rapid growth, there are many frameworks being developed; however, there is no universally adopted framework that captures the AI/ML development lifecycle and common architectures. Although, there are several options available that may be used by MDMs and HDOs to help more systematically understand the depth and breadth of considerations for cybersecurity, safety, and other concerns, including those developed by NIST [11], FDA [12], Organization for Economic Co-operation and Development [13], and Berryville Institute of Machine Learning [14].

Generally, such frameworks include lifecycle phases such as:

- Raw data collection and conversion, which could include human-generated or machine-generated data.
- Model building, training, and tuning.
- Testing.
- Verification and validation.
- Deployment and operation.
- Maintenance/update.

Another phase that is gaining increasing attention—end-of-life (EOL) or end-of-service (EOS)—was not a subject of any of our interviews and does not appear to be covered in any detail in commonly used documents that contain their own AI/ML lifecycles.

With AI/ML being heavily data-driven, there are typically different kinds of data that are generated, managed, and secured, such as:

- Raw data, which is the original source of data from which the system learns. This data is often unstructured and not explicitly labeled.

[5] This paper uses the following definitions: Artificial Intelligence (AI) is defined as "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments" [21] and Machine Learning (ML) is defined as "a set of techniques that can be used to train AI algorithms to improve performance at a task based on data" [16].

- Training datasets, which are often more structured and include labeled examples of key concepts, enabling effective processing by AI/ML models.
- Testing, validation, or evaluation datasets, which are used to assess model performance, including generalization to unseen data and detection of issues such as model overfitting.
- Adversarial testing datasets, used to evaluate model robustness against inputs designed to cause misclassification, unexpected behavior, or security vulnerabilities.
- Models, which are often large, complex data structures, including learned parameters such as weights and associated training hyperparameters that influence the model's behavior.

Depending on the technology in use, AI/ML system architectures may include the following components:

- A component that collects and possibly transforms[6] data so that it can be fed into learning components.
- Learning algorithms (i.e., computational methods that learn patterns in data, such as classification or cluster analysis).
- Inference algorithms (i.e., computational methods that make predictions or inferences using probabilistic models as part of the input, such as Bayesian inference or genetic algorithms).
- Sub-systems like Natural Language Processing (NLP) components.
- Model management, including MDM-controlled construction or import from third-party sources and maintenance of "inventory."
- Components that obtain new data and integrate it into existing data for dynamic training to improve the system's performance and accuracy.

AI/ML can be broken down into two broad capabilities:

- "Discriminative" (e.g., predicting outcomes or performing classification based on input data).
- "Generative" (e.g., producing or summarizing outputs).

Discriminative AI/ML techniques have been used for many years and are regarded as relatively mature (e.g., diagnostics in radiology). However, while there has been significant interest in generative technology such as Large Language Models (LLMs), their use within medical devices has not been widely adopted.

While AI/ML is an area of interest for many MDMs, they vary in how they use AI/ML in two key areas:

- How independently the AI/ML operates (i.e., whether its use is fully automated versus having a human in the loop).
- The types of tasks for which the AI/ML is used (e.g., diagnosis, decision support, report writing for clinicians, etc.) Due to patient safety considerations, MDMs and HDOs are very careful about integrating AI/ML components in roles that directly involve patient treatment.

[6] Data transformation may be described by terms such as "data cleaning," "data preprocessing," or "data wrangling."

AI/ML technologies, algorithms, and associated data can vary widely in their maturity in terms of how well their strengths and limitations are understood, and where they may fit within medical devices. AI/ML-focused tooling is available to detect potential problems in terms of safety, security, and correctness; capabilities of protection mechanisms that limit the scope of AI/ML-oriented vulnerabilities and exploitation; etc. In addition, MDMs, HDOs, and other organizations may vary widely in their adoption of AI/ML due to factors such as their risk tolerance, the flexibility of their existing processes to integrate such technologies into current products versus new ones, etc.

AI/ML adoption varies in different types of medical devices, although our interviews and follow-up analyses did not investigate possible reasons for this variation.

Dating back to November 1995, AI/ML technologies have been used in numerous medical devices approved by FDA. For example, in an FDA-maintained database as of September 30, 2025, there are 1,357 entries listed, with 1,039 devices related to Radiology (76.56%) and 130 devices related to Cardiovascular (9.58%) [15].

AI/ML software can be used in different contexts, such as but not limited to:

- Acting as a medical device itself (Software as a Medical Device) [16].
- As a component of a medical device that directly contributes to the device's essential performance,[7] such as analysis and interpretation of an MRI image, calculation of bolus dosage in an insulin pump, and/or connectivity with related systems.
- As support for the use of medical devices, such as healthcare dashboards or analytics that use AI/ML to improve decision making, learn patterns in healthcare data, assist with triage, and improve patient outcomes.
- To support clinical operations that are not directly tied to medical devices themselves, such as use of LLMs to write diagnostic summaries or visit reports.
- To generate other software that is used in medical devices.

Note that even if a medical device does not contain AI/ML capabilities, it is possible that the MDM or its third-party developers use AI-generated code. Stakeholder interviews did not include discussion of the use of AI-generated code by the MDM or whether they ask their third-party suppliers about AI usage, so it is unclear how common this practice is.

## 4.2 Threats and Risks

Incorporating AI/ML algorithms into medical devices may introduce opportunities for novel cyberattacks that can impact patient safety by producing outputs that lead to misdiagnoses or under/over treatment. As with cloud computing services, AI/ML algorithms can be used throughout the lifecycle of a medical device, during design and development, operational deployment of essential functionality, or for providing ancillary services.

[7] The definition of "essential performance" can be found in IEC 60601-1 [22].

Even if AI/ML is not incorporated into a medical device, the device's software may have been designed and developed with the help of an AI/ML system. It is possible that AI-generated code could introduce unusual bugs that are not typically encountered in human-written code, which could make them more difficult to diagnose (especially due to the black box nature of many AI/ML algorithms) and/or more difficult to even detect in the first place, since many software security analysis capabilities are focused on established, well-defined code patterns.

For medical devices that use AI/ML to perform their essential functionality or provide ancillary services, there are threats and risks specific to AI/ML technology. Some threats are exploited during algorithm development (e.g., altering training data), others during operational use (e.g., prompt injections or adversarial inputs[8]. Some cyberattacks achieve goals that may directly impact patient safety and privacy, while others are used to gain insight into the AI/ML model in order to develop further attacks to achieve the adversary's real objectives. Privacy is also a consideration by MDMs for AI/ML, due to Health Insurance Portability and Accountability Act (HIPAA[9]). Common attacks can leak information that can violate organizations' intended privacy policies, such as membership inference attacks that may be able to determine if individuals were part of the original training data. Similar kinds of attacks could be successful in extracting other critical, confidential information that is not directly related to privacy, such as API keys.

AI/ML technology is heavily dependent on the quality and integrity of its data. If an adversary can alter or "poison" the data throughout any stage of the AI/ML lifecycle, or otherwise influence how the component analyzes the data, then the data may not be sufficiently trustworthy, creating opportunities for the adversary to compromise the system. As a result, diligence to ensure the integrity and authenticity of all data across the AI/ML lifecycle is an important consideration. This includes but is not limited to: raw data; training data (whether "labeled" or "unlabeled"); data that is used for verification, validation, or other testing; the models themselves (including weights and input parameters); behavior-influencing inputs such as prompts for LLMs, which effectively act as code but may be subject to manipulation; data for sub-systems (e.g., natural language models); etc.

Finally, AI/ML is primarily implemented and managed using software, so it is subject to the same weaknesses (coding errors or design flaws) that can occur in all software. As a result, common security vulnerabilities or safety-relevant errors may be introduced that are not directly related to the AI/ML behavior itself. Malicious attackers can exploit these vulnerabilities to affect the availability, integrity, and confidentiality of the AI/ML-enabled medical device.

[8] Adversarial inputs are specially crafted data used to produce incorrect output from a classification system. Prompt injections are malicious prompts to an LLM to cause it to behave in unintended ways. See MITRE ATLAS [27] for more details on AI/ML specific threats.

[9] https://www.hhs.gov/hipaa/for-professionals/index.html

## 4.3 Challenges

There are several challenges that MDMs face as they consider adoption of AI/ML capabilities and how to secure them.

**Unpredictable or difficult-to-control behavior.** Many AI/ML capabilities are "stochastic" in nature (i.e., they inherently exhibit variability in their outputs, even when presented with identical inputs). This non-deterministic, unpredictable behavior contrasts with traditional deterministic programming, where outputs are predictable and repeatable as formally enforced by the restrictions of the programming language and environment. For traditional programming, there are well-established methods for ensuring that code does not behave outside of expectations, and the code can be audited for potential errors or security vulnerabilities. With AI/ML, the resulting behavior will be less predictable and repeatable, which may have potentially severe implications for safety and correctness. For example, generative AI applications such as LLMs can produce "hallucinations," which are outputs that are false, erroneous, or nonsensical, even if they are plausible outputs from the model itself. Depending on the context in which the AI/ML functionality is used, such "hallucinations" can affect the device's behavior or outputs in undesirable ways that affect patient safety, such as misdiagnosing or failing to recognize changes in patient conditions. In some cases, attackers can increase the chances or severity of hallucinations. Alternately, attackers could use various injection techniques such as prompt injection, to influence what commands are provided to the AI/ML component and/or executed by the system.

**Limitations of traditional defenses and uncertain success of new defenses.** Traditional software security techniques often assume clearly defined, predictable behavior based on the same set of inputs, which can increase confidence in effectiveness of software security techniques with well-defined testing and formal or semi-formal analysis of code artifacts, such as source code. Such defenses may still be appropriate for AI/ML because it is still managed by and implemented using software, but there are important stumbling blocks. With traditional software, most errors can be detected in a repeatable fashion by re-running the code in the same environment using the same inputs. More importantly, debugging and other code analysis mechanisms are available that provide traceability into the software, such as how the code executes step-by-step; which variables are consulted; a full "stack trace" of called functions; etc. With AI/ML, however, the underlying behaviors are typically a "black box" in which results can vary with each run, and there is not a good way to perform step-by-step analysis or otherwise inspect the logic and data flow to get consistent results or to understand why a particular decision was made. That is, there is insufficient information to have a complete, comprehensive understanding of why a particular error occurred, or to closely analyze the code to gain confidence that it will perform within expectations given particular inputs. With AI/ML, the result is a disconnect in which a problem arises, but there is not a detailed understanding of how the problem occurred, and there cannot be complete confidence that a solution will work 100% of the time. Various technologies can help to better control AI/ML behavior, such as "guardrails" that watch for unexpected, potentially malicious inputs and/or behavior, or Retrieval Augmented Generation (RAG) to minimize hallucinations and maintain focus on domain-specific information. However, these defenses are new and rapidly evolving. There might not be solid guarantees or easily available metrics in terms of success in preventing unexpected behavior and efficiency (in terms of speed or resource consumption, e.g., latency or cost). For example, many

"jailbreak" attacks are targeted at systems that are already attempting to perform some kind of restriction of inputs, behavior, or output. In general, the unpredictable nature of AI/ML is a major concern for MDMs, especially in applications that can lead to significant adverse impact on patients, so AI/ML is only adopted after careful risk analysis.

**Locked vs. adaptive mode.** MDMs can plan how and when learning is performed in their AI/ML components. In adaptive mode, new data is regularly learned and integrated into the system during operation, allowing for ongoing improvements on a regular or continuous basis. In locked mode, learning may be performed offline during development, with more distinctly separate phases of model training, testing, etc.

- **For adaptive mode:** Customization could happen more quickly, but this could omit important validation and verification steps in which the system could independently verify or validate the model against a fixed test dataset, potentially missing indications of a cybersecurity attack. Adaptive mode may also further expose the system to attacks using poisoned or malicious data. There is also a risk that AI-generated results are fed back into the model, thus degrading model performance in comparison to the real-world data that may have been used in the original creation of the model[10] (e.g., when applied to patient subgroup data that is different from the datasets on which the model was trained). There are also important availability considerations for how to interface with online services from AI/ML providers, which may go offline or otherwise not be reachable upon demand.
- **For locked mode:** Once learning has been completed, the component could be deployed in a more predictable way, offering notions of "versioning" like that of software, and providing clear opportunities to evaluate performance against a fixed test dataset. However, the process may be more labor-intensive, or the AI/ML capability may not be updated as quickly compared to adaptive mode.

**Relation with cloud technologies.** Some AI/ML capabilities may operate in the cloud due to insufficient computing resources in the edge computing components themselves (e.g., small wearable or implanted medical devices). As a result, the MDM inherits the potential risks and challenges of adoption of cloud computing, including the potential for loss of cloud availability, (i.e., the inability for the medical device to perform correctly if the cloud service becomes unavailable due to a direct cybersecurity attack or an inadvertent outage originating from the provider's own errors; see Section 3.3). While cloud technology may have been adopted by MDMs for other reasons, the cybersecurity risk analysis for an AI/ML-enabled medical device independently considers cloud dependencies.

**Clinical usefulness and acceptance.** While there is high interest in AI/ML by MDMs, some MDMs noted in our interviews that hospitals and clinicians may struggle to clearly understand the benefit in AI/ML solutions, which may slow adoption, especially in areas that could affect patient safety. Even if clinical use cases with high potential are identified, they might not be accepted by clinicians themselves. Factors of concern may include unpredictable behaviors; requiring labor-intensive, hands-on interaction; the amount of resource-intensive computing needed; etc.

[10] This performance degradation is described by terms such as "model drift," "model collapse," etc. It may have different causes [26].

## 4.4 Mitigations

MDMs may consider the following high-level mitigations for the risk of integrating AI/ML into their medical devices.

**A secure learning environment.** An important consideration is for all training data, parameters, label data (if any), models, related code, and associated processing systems to be secured properly from potential threats, including prompts and other text-based "code." This security helps to ensure that the test dataset is separated from the training dataset in order to avoid data contamination (i.e., ML data poisoning). Numerous resources exist that provide more technical details, including those mentioned in Section 4.1, although as of this writing, they are rarely specific to healthcare.

**Guardrails with sufficient robustness testing.** Unlike traditional coding, AI/ML behavior is often unpredictable, so its behavior might change even when given the same set of inputs (i.e., the behavior may be "stochastic" [17]). MDMs can consider appropriate guardrails or other protection mechanisms to reduce the risk of AI/ML behavior operating outside of expectations. However, guardrails cannot necessarily guarantee full control over behavior, and there is extensive research into "jailbreak" techniques, some of which may be very narrowly scoped to the domain in which the AI/ML is operating. Depending on the underlying AI/ML technology being used, other approaches can be considered to control behavior, such as RAG, prompt engineering, and careful configuration of "hyperparameters" that influence the range of possible model behavior, such as LLM settings for temperature, Top-P, and Top-K. For cybersecurity, a targeted subset of data permutations can be integrated into the testing process to find more obvious problems. Manual red teaming can be effective as well, since techniques may vary depending on the technologies involved, and Subject Matter Experts (SMEs) can explore and exploit idiosyncratic behaviors of the system in ways that cannot be automated.

**AI/ML security as a part of the overall software security program.** As with traditional software components, threat modeling clearly identifies and closely analyzes AI/ML components.[11] As with other components in the medical device, design could consider controls around interactions between components, especially when data or transactions across components cross a trust boundary (e.g., with cloud-based AI/ML components, or using data or models provided by third parties). Following the "Least Privilege" principle[12] can help ensure that AI-enabled agents or subsystems do not have more privileged access into the rest of the system (or connected systems) than they should. Integrity of outcomes is also important, especially since AI/ML outputs can be unpredictable and will not necessarily follow a specification or well-defined range of expected outputs, as usually occurs with traditional programming.

**Risk and liability analysis during acquisition.** During acquisition of components involving AI/ML, diligent analysis of potential risks, including cybersecurity risk is an important consideration. Risk considerations may also include potential liability concerns if the AI/ML component fails in ways that could affect patient safety, confidentiality, product correctness, etc.

[11] FDA's "Cybersecurity in Medical Devices: Quality System Considerations and Content of Premarket Submissions" [19] discusses FDA's recommendations for threat modeling in general and MITRE's threat modeling playbook [9] offers some recommendations on implementing threat modeling. MITRE has developed ATLAS, a framework for assessing threats specific to AI/ML [27].

[12] "The principle that a security architecture is designed so that each entity is granted the minimum system authorizations and resources needed to perform its function" [28].

# 5 Post Quantum Cryptography (PQC)

## 5.1 Overview

### 5.1.1 Background

Cryptography is a fundamental technique to secure information systems of all kinds, including medical devices. Cryptographic algorithms are roughly split into two major categories: symmetric algorithms (e.g., block ciphers, hash functions, and their applications) and asymmetric algorithms, also known as public-key cryptography.

Since Peter Shor's landmark 1994 paper "*Algorithms for quantum computation: Discrete logarithms and factoring*" [18], we have known that the mathematical problems underlying all historically popular, widely deployed forms of public-key cryptography[13] would be solvable efficiently on a quantum computer. As such, a scalable quantum computer (Cryptanalytically Relevant Quantum Computer" (CRQC)) with sufficient power and sophistication that can run Shor's algorithm would pose a threat to any security properties enforced by those algorithms, including those in existing controls in medical devices.

Shor's paper sparked interest in developing quantum computers, which were a purely theoretical idea at the time,[14] and research in the area has been described as a global technological race towards scalable quantum computing with an emphasis on mitigating risks to vulnerable cryptographic systems.[15]

Given that medical devices may be used past their end-of-life and end-of-support dates, cybersecurity controls used to originally design medical devices are likely to become vulnerable once this algorithm is broken. Continuing to protect such information with vulnerable controls may allow an attacker with CRQC access to exfiltrate encrypted information from the device (such as patient data, intellectual property, and device telemetry), and decrypt the data, and gain access to still-relevant confidential information. This is known as the "harvest-now, decrypt-later" threat.

Because of the significance of the quantum threat, especially given the harvest-now decrypt-later approach, the National Institute of Standards and Technology (NIST) initiated its "post-quantum process" in 2016 to mature and eventually standardize new asymmetric algorithms specifically evaluated on the basis of their security, both against quantum and classical attacks.[16] These algorithms, known as "post-quantum cryptography," are a conventional (non-quantum) cryptographic defense against the quantum threat.

[13] Including in particular the RSA algorithm, the Diffie-Hellman key exchange protocol, and Elliptic Curve Cryptography.

[14] https://quantumzeitgeist.com/a-brief-history-of-the-quantum-computer/

[15] https://www.marketsandmarkets.com/Market-Reports/quantum-computing-market-144888301.html, https://thehill.com/opinion/technology/4642324-the-quantum-computing-race-is-on/, https://www.forbes.com/sites/drektadang/2025/03/09/recent-breakthroughs-accelerate-the-race-for-quantum-computing/

[16] https://csrc.nist.gov/projects/post-quantum-cryptography

In recognition of the potential threat, the U.S. government has introduced a number of policies through executive orders, national security memoranda, and legislation, encouraging the transition to post-quantum cryptography within the government, both in national security/defense applications and in Federal Civilian Executive Branch (FCEB) agencies. See Appendix A-4 for a fuller presentation of the variety and scope of official government policies on the issue.

### 5.1.2 Threats

This section will focus on potential post-quantum threats to medical devices and their systems, and the health sector.

The threats are discussed in the context of their impact on patient safety and the healthcare system overall.

Cryptographic algorithms are one of the essential building blocks ensuring the cybersecurity of medical devices and related systems, which is integral to maintain their functionality and ensure patient safety [19]. We discuss some examples of specific impacts below, but the lack of effective cryptography in these systems, if an attacker has access to a CRQC and these systems are still using CRQC-vulnerable cryptography, would be impossible to address in a simple patch or update: as they would be using an insecure design that cannot be easily repaired.

Digital signature algorithms, a form of asymmetric cryptography, are electronic methods designed for computers to both create and check. They are primarily used for electronic authentication and non-repudiation. This form of authentication is used frequently in all types of controls. If circumvented, that control breaks down and the safety of patients may be impacted.

Safety impacts from a quantum attack on CRQC-vulnerable encryption or key exchange would mean unapproved access to devices and encrypted information and may lead to patient harm. This includes unapproved access to information such as passcode, pin information, etc., that can be used to gain access to medical devices, systems, or facilities.

Medical devices using currently acceptable cryptographic controls may be vulnerable to quantum attacks. For example, the following capabilities that use current acceptable security controls will no longer be secured once CRQC is broken:

- Securing unapproved access to controlled substances.
- Securing unapproved access to high-risk medical devices.
- Falsifying software updates, resulting in impacts ranging from availability issues to complete attacker control of those devices and systems.
- Forgery of identification credentials, which can lead to control of medical devices and their systems.
- Accessing or altering sensitive information about a patient that could be used to harm them, for instance, information that can impact medical diagnosis.
- Interfering with the operations of the healthcare delivery organization, indirectly impacting patient safety.

- Accessing proprietary communication about potential vulnerabilities in a facility and then exploiting them in a dangerous way.
- Accessing proprietary information about a medical device or facility, which could be analyzed for a design flaw and exploited in a way that impacts safety.

## 5.2 PQC Transition and Implementation

Medical device manufacturers and healthcare industry may want to plan their transition to utilizing PQC to address the threat posed by quantum technology.

PQC algorithms exist both for confidentiality applications (encryption) and for authentication and non-repudiation applications (digital signatures), which can be used as a substitute for existing vulnerable algorithms.

However, it is important to note that transitioning from the widely used vulnerable existing cryptography to PQC is not a simple proposition and will take time and planning. In this section, we provide a rough overview of the space and common approaches to PQC migration.

### 5.2.1 Organizational Goals

Given the interconnected nature of medical devices and healthcare systems, it could be useful for MDMs and Healthcare Organizations to coordinate the PQC migration. An additional consideration for PQC is to individually address challenges within devices themselves and through the clinical ecosystem. There are two major categories of goals that an organization may have in the PQC space:

- To ensure that their devices are cybersecure, it will be important for MDMs to address PQC threats in devices and systems they build and maintain.[17]
- It will also be important for MDMs and Healthcare Organizations to address PQC to protect themselves and their operations from quantum threats. This goal is supported by their understanding of the range of cryptographic systems they use and planning how to adopt PQC and eventually phase out legacy algorithms while maintaining their ability to deliver healthcare safely and effectively.

The types of activities involved in achieving those goals are significantly different.

- MDMs, addressing the quantum threat in medical devices and systems they build, secure them against the quantum threat, addressing all technical concerns and considerations.
- Developing an overall strategy and taking responsibility for decisions across a wide range of systems and products, including those that MDMs and Healthcare Organizations do not develop or control, to protect their patients and their operations.

[17] FDA's premarket guidance [19] provides recommendations on the selection and implementation of cryptographic algorithms and protocols used by medical devices.

### 5.2.2 Developing a Strategic Plan

To proactively assess an organization's exposure to the quantum threat, including the devices they manufacture and any products they rely on, a strategic plan to address these threats can be developed.

The strategic plan could include four major categories:

- **Goal setting** – Identify goals and timelines to phase out existing insecure controls and implement PQC.
- **Gathering information** – Assess the products and systems. (1) Identify which products and systems use quantum-vulnerable asymmetric cryptography controls, (2) prioritize the products and systems as part of a phased plan, and (3) identify the available relevant PQC controls.
- **Resources and implementation** – Identify and allocate resources, include verification and validation protocols for the chosen PQC controls, and develop training in preparation for the implementation of the chosen controls. The goals are re-evaluated throughout the process to ensure their applicability. Adaptation to the goals is assessed, and the verification and validation is likely to be updated accordingly.
- **Automated cryptographic discovery and inventory (ACDI) tools** have been recognized as helpful in this area. Advancing capabilities have been a goal of the National Cybersecurity Centers of Excellence project on Migration to Post-Quantum Cryptography.[18] However, tools of this nature tend to be focused on discovery in the general enterprise information technology domain; they do not claim to be helpful in assessing vulnerabilities in special purpose equipment such as medical devices or specialized medical systems.

General enterprise information technology organizations acted together to set industry-wide goals, signaling as a sector that products that will help get them to their goals were needed. This is already having an effect in general enterprise information technology, but where medical organizations rely on specialized products and systems that are not used broadly (say) in the defense sector or in federal government organizations, the demand signal to manufacturers may be a challenge.

### 5.2.3 PQC Transition and Implementation Considerations

**Technical requirements.** In addition to addressing the quantum threats in their devices, there are technical challenges that post-quantum controls may present. The new post-quantum algorithms are not vastly different in terms of overall computation time, but they may require more memory, more code, and longer messages, and correspondingly may require more power and may incur additional expenses due to the components needed.

**Interoperability with legacy devices**. Many medical devices interoperate with other medical devices and systems. It is likely that new devices with PQC controls will interoperate with existing systems or medical devices that do not implement post-quantum cryptography controls. As long as legacy systems exist, the quantum risk is not eliminated. This interface with legacy devices remains a threat to assess and address.

[18] https://www.nccoe.nist.gov/sites/default/files/2022-07/pqc-migration-project-description-final.pdf, https://www.nccoe.nist.gov/sites/default/files/2023-12/pqc-migration-nist-sp-1800-38b-preliminary-draft.pdf

This could involve handling multiple types of cryptography in at least one endpoint of any cryptographic communication, which would further increase necessary system complexity for those endpoints.

**Transition timelines.** The time it may take to gradually phase out or update devices that are using traditional quantum vulnerable asymmetric cryptography controls will vary based on many factors, including but not limited to: engineering considerations for the devices and systems themselves, financial factors, legal mandates, interoperability and continuity plans, device lifetime in the field, feasibility of updating existing devices vs. replacing them, and the overall environment of use and other related and safety considerations. For example, if an implantable medical device cannot be reprogrammed without physical access to the device, it would take a medical procedure to remove or alter the cryptography. This is unlikely to occur solely to adjust the cryptography. Medical procedures incur risk and may be considered appropriate only when they overall are in the best interest of the patient.

# 6 Summary

Emerging and evolving technologies promise advances in patient care, but at the same time introduce new cybersecurity risks. Managing these risks doesn't necessitate an entirely new approach but builds upon existing practices. Safeguarding medical devices against the threat of quantum cryptographic attacks and cyberattacks against cloud infrastructure and AI/ML algorithms used by the devices can be achieved by considering how to incrementally adapt processes, such as using SBOMs to manage vulnerabilities and creating threat models, to include cloud, AI/ML, and cryptographic components.

Governance is central to managing cybersecurity risk. The frameworks and architectures lead to rethinking roles and responsibilities. When acquiring systems with these new technologies, contracting language and Cloud SLAs can clarify the cybersecurity expectations of the various parties. Planning and roadmaps for adoption and migration can be developed to ensure a smooth transition.

The only constant is change: technology will continue to evolve. It is important to develop medical devices that can be updated so they can continue to provide needed care without becoming vulnerable and introducing threats to patients and the environments in which they are used.

# Glossary

| Acronym | Definition |
|---|---|
| AI | Artificial Intelligence |
| CAVEaT | Cloud Adversarial Vectors, Exploits, and Threats |
| CI/CD | Continuous Integration/Continuous Delivery |
| CISA | Cybersecurity and Infrastructure Security Agency |
| CRQC | Cryptanalytically Relevant Quantum Computer |
| FDA | Food and Drug Administration |
| FFRDC | Federally Funded Research and Development Center |
| HDO | Healthcare Delivery Organization |
| HIPAA | Health Insurance Portability and Accountability Act |
| IaaS | Infrastructure as a Service |
| LLM | Large Language Model |
| MDM | Medical Device Manufacturer |
| ML | Machine Learning |
| NSA | National Security Agency |
| NIST | National Institute of Standards and Technology |
| OECD | Organization for Economic Co-operation and Development |
| PaaS | Platform as a Service |
| PQC | Post-Quantum Cryptography |
| SaaS | Software as a Service |
| SBOM | Software Bill of Materials |
| SLA | Service Level Agreement |

# Appendix A - Resources

## A-1 General

This section collects some resources applicable to general cybersecurity considerations in medical devices.

- FDA Cybersecurity Guidance:
  - Cybersecurity in Medical Devices: Quality System Considerations and Content of Premarket Submissions | FDA.
  - Postmarket Management of Cybersecurity in Medical Devices | FDA.
- Threat Modeling:
  - Playbook for Threat Modeling Medical Devices | MITRE – MITRE's playbook on applying the four question threat modeling framework to medical devices.
- SBOM:
  - Software Bill of Materials (SBOM) | CISA – CISA's SBOM site collects various resources on creating and using SBOMs, tooling, and sharing.
  - Data Normalization Challenges and Mitigations in Software Bill of Materials Processing | MITRE – MITRE white paper on data normalization challenges in processing SBOMs.

## A-2 Cloud Computing

This section collects some of the key resources on cloud computing, including general and cybersecurity-specific frameworks.

- NIST Cloud Computing Framework:
  - NIST SP 800-145, The NIST Definition of Cloud Computing.
  - NIST Special Publication 800-146, Cloud Computing Synopsis and Recommendations.
- CISA Resources:
  - CISA's Fact Sheet on free tools for securing cloud environments and other resources: Free Tools for Cloud Environments.
- Threat-Informed Defense:
  - Matrix - Enterprise - Cloud | MITRE ATT&CK® – The ATT&CK Cloud Matrix includes SaaS and IaaS platforms.
  - Home - Mappings Explorer – The Mappings Explorer maps security controls and capabilities to the adversary behaviors catalogued in MITRE ATT&CK. It includes Mapping Frameworks for Azure, Google Cloud Platform, and Amazon Web Services.
  - CAVEaT™ | CSA – MITRE and the Cloud Security Alliance are developing a cloud-specific threat model.

- Cloud Computing Best Practices:
  - Cybersecurity and Infrastructure Agency, United States Digital Service, and Federal Risk and Authorization Management Program Cloud Security Technical Reference Architecture v.2.
  - NSA and CISA Cybersecurity Information Sheet on defending CI/CD environment - CSI_DEFENDING_CI_CD_ENVIRONMENTS.PDF.
  - Introduction - OWASP Cheat Sheet Series – OWASP includes several cheat sheets relevant to securing cloud computing, including Docker Security, Kubernetes Security, Secrets Management, and Secure Cloud Architecture.
- Cloud Service Provider Cybersecurity Information:
  - Amazon Web Services (AWS):
    - AWS Security Reference Architecture (AWS SRA) – AWS Prescriptive Guidance.
    - Penetration Testing – Amazon Web Services (AWS).
  - Google Cloud Platform (GCP):
    - Cloud Security Best Practices Center | Google Cloud.
    - Cloud Security FAQ - Google Cloud Platform Console Help (Question on penetration).
  - Microsoft Azure:
    - Security best practices and patterns – Microsoft Azure | Microsoft Learn.
    - Penetration testing | Microsoft Learn.

## A-3 AI/ML

This section collects some key resources on securing AI/ML systems.

- General:
  - Artificial Intelligence in Software as a Medical Device | FDA – FDA website collects information on regulatory considerations (U.S. and global) on AI-enabled medical devices.
  - OECD Framework for the Classification of AI systems | OECD – OECD's framework for characterizing AI systems from a policy perspective.
- Risk Assessment and Threat Modeling:
  - Berryville Institute of Machine Learning – Architectural risk analysis framework of generic and LLM machine learning systems.
  - NIST AI RMF Resources - AIRC – NIST's AI Risk Management Framework, Playbook with suggestions for achieving the AI RMF outcomes, and other resources.
  - MITRE ATLAS™ – MITRE's Adversarial Threat Landscape for Artificial-Intelligence Systems (ATLAS) is a knowledge base of adversary tactics and techniques against AI-enabled systems, that complements ATT&CK.
- Best Practices for Developing and Operating Machine Learning Models:

- Good Machine Learning Practice for Medical Device Development: Guiding Principles | FDA – FDA and UK Medicines and Healthcare products Regulatory Agency's guiding principles to inform the development of medical devices that use AI/ML.
- Secure AI Model Ops - OWASP Cheat Sheet Series – OWASP's guide for operating and deploying AI/ML systems.

## A-4 PQC

In recognition of the potential threat to digital systems of all kinds from the development of a CRQC, the U.S. government has introduced numerous initiatives and laws for the transition to post-quantum cryptography to mitigate the quantum threat.

- In June of 2025, the President issued Executive Order 14144 "Strengthening and Promoting Innovation in the Nation's Cybersecurity,"[19] which was amended in Executive Order 14306, "Sustaining Select Efforts to Strengthen the Nation's Cybersecurity and Amending Executive Order 13694 and Executive Order 14144."[20] As amended, this order places some additional requirements on the government: to create and maintain a list of product categories in which products supporting PQC are widely available by December 1, 2025, and to issue requirements on the government to support TLS 1.3 by January 2, 2030.[21]
- NIST issued the first final general purpose post-quantum cryptographic algorithm standards (FIPS 203, 204, and 205) in August of 2024.[22] NIST has since categorized these new standards as "Acceptable" algorithms and has announced plans to categorize all CRQC-vulnerable asymmetric algorithms as "Disallowed" of 2035—in line with the NSM-10 deadline.[23] This categorization affects what algorithms may be officially certified in products in the Cryptographic Module Validation Program (CMVP), a certification process for many government applications and in various industries as voluntary best practices.
- H.R.7535 The Quantum Computer Cybersecurity Preparedness Act[24] was signed into law in December of 2022, following NSM-8 and NSM-10.

[19] https://www.federalregister.gov/documents/2025/01/17/2025-01470/strengthening-and-promoting-innovation-in-the-nations-cybersecurity

[20] https://www.whitehouse.gov/presidential-actions/2025/06/sustaining-select-efforts-to-strengthen-the-nations-cybersecurity-and-amending-executive-order-13694-and-executive-order-14144/

[21] This requirement concerning TLS 1.3 is PQC-related in the sense that TLS 1.3 – the latest version of TLS – is the only version that will be updated to incorporate PQC algorithms; see https://datatracker.ietf.org/doc/draft-ietf-tls-tls12-frozen/

[22] https://www.nist.gov/news-events/news/2024/08/nist-releases-first-3-finalized-post-quantum-encryption-standards

[23] The initial public draft of "Transition to Post-Quantum Cryptography Standards (NIST IR-8547)," published in November 2024 (https://nvlpubs.nist.gov/nistpubs/ir/2024/NIST.IR.8547.ipd.pdf), provides timelines for transition.

[24] https://www.congress.gov/bill/117th-congress/house-bill/7535/text, Became Public Law 117-260: https://www.congress.gov/117/plaws/publ260/PLAW-117publ260.pdf

- The National Security Agency (NSA) issued their Commercial National Security Algorithm suite version 2.0 (known as CNSA 2.0) in September of 2022.[25] This outlines an updated set of commercial (public) algorithms that will eventually be included in national security applications. All asymmetric algorithms in the suite are post-quantum algorithms. The initial stated goal of NSA is to have national security applications exclusively using CNSA 2.0 by 2033, although this goal has since been updated to December 31, 2031.[26]
- National Security Memorandum 10 ("National Security Memorandum on Promoting United States Leadership in Quantum Computing While Mitigating Risks to Vulnerable Cryptographic Systems") was issued in May of 2022, and mandates that all of the U.S. federal government "prioritize the timely … transition of cryptographic systems to quantum-resistant cryptography."[27]
- National Security Memorandum 8 ("Memorandum on Improving the Cybersecurity of National Security, Department of Defense, and Intelligence Community Systems") was issued in January of 2022, and required the NSA to articulate post-quantum modernization goals and issue timelines.[28]

H.R.6227 – National Quantum Initiative Act was signed into law in December of 2018.[29] The act calls for the continued leadership of the United States in Quantum Information Science (QIS) and its applications, and for a coordinated federal program to accelerate quantum research and development for the economic and national security of the United States.

[25] https://media.defense.gov/2025/May/30/2003728741/-1/-1/0/CSA_CNSA_2.0_ALGORITHMS.PDF

[26] In addition to the CNSA 2.0 suite of asymmetric algorithms, NSA has articulated a desire to use AES-256, a symmetric algorithm with larger key sizes than the normal 128 bits used in common practice. However, this recommendation is unique to CNSA 2.0 and is not raised as a concern in NSM-10 or Public Law 117-260, and NIST has announced no plans modify its recommendations regarding symmetric cipher key lengths, despite addressing the issue of security strength generally as recently as fall of 2024.

[27] https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2022/05/04/national-security-memorandum-on-promoting-united-states-leadership-in-quantum-computing-while-mitigating-risks-to-vulnerable-cryptographic-systems/. See also M-23-02, "MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES," directing agencies to answer a specific data call regarding PQC inventory and cost planning in support of NSM-10 implementation.

[28] https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2022/01/19/memorandum-on-improving-the-cybersecurity-of-national-security-department-of-defense-and-intelligence-community-systems/

[29] https://www.congress.gov/bill/115th-congress/house-bill/6227 became Public Law 115-368.

# References


[1] IMDRF, "Principles and Practices of Cybersecurity for Legacy Medical Devices (IMDRF/Cyber WG/N70Final:2023," 11 April 2023. [Online]. Available: https://www.imdrf.org/documents/principles-and-practices-cybersecurity-legacy-medical-device.

[2] P. Mell and T. Grance, "The NIST Definition of Cloud Computing (NIST SP 800-145)," September 2011. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf.

[3] E. Udoh, B. Patterson and S. Cordle, "A Performance Analysis of Cloud Computing Using the Balanced Scorecard Approach)," in 2014 Annual Global Online Conference on Information and Computer Technology, Louisville, KY, 2014.

[4] J. Davis, "Elekta sued over ransomware attack; Intermountain, Advocate Aurora added to breach tally," 20 July 2021. [Online]. Available: https://www.scworld.com/analysis/elekta-sued-over-ransomware-attack-intermountain-advocate-aurora-added-to-breach-tally.

[5] J. Terra, "What is a Cloud Service Level Agreement?: A Beginner's Guide," 14 August 2024. [Online]. Available: https://pg-p.ctme.caltech.edu/blog/cloud-computing/what-is-a-cloud-service-level-agreement.

[6] National Institute of Standards and Technology, "Cloud Computing Synopsis and Recommendations (NIST SP 800-146)," May 2012. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-146.pdf.

[7] Healthcare and Public Health Sector Coordinating Councils, "Health Industry Cybersecurity -Model Contract-language for Medtech Cybersecurity," March 2022. [Online]. Available: https://healthsectorcouncil.org/wp-content/uploads/2022/05/HSCC-Model-Contract-language-for-Medtech-Cybersecurity-2022.pdf.

[8] Cybersecurity and Infrastructure Security Agency, United States Digital Service, and Federal Risk and Authorization Management Program, "Cloud Security Technical Reference Architecture, version 2," June 2022. [Online]. Available: https://www.cisa.gov/sites/default/files/2023-02/cloud_security_technical_reference_architecture_2.pdf.

[9] The MITRE Corporation, "Playbook for Threat Modeling Medical Devices," 30 November 2021. [Online]. Available: https://www.mitre.org/sites/default/files/2021-11/Playbook-for-Threat-Modeling-Medical-Devices.pdf.

[10] MITRE and Cloud Security Alliance CAVEaT Working Group, "Cloud Adversarial Vectors, Exploits, and Threats (CAVEaT): An Emerging Threat Matrix for Industry Collaboration," 2021. [Online]. Available: https://cloudsecurityalliance.org/research/working-groups/caveat.

[11] National Institute of Standrads and Technology, "Artificial Intelligence Risk Management Framework (AI RMF 1.0) - NIST.AI.100-1," January 2023. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf.

[12] U.S. Food and Drug Administration, "Good Machine Learning Practice for Medical Device Development: Guiding Principles," 25 March 2025. [Online]. Available: https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles.

[13] Organisation for Economic Co-operation and Development, "OECD Framework for the Classification of AI Systems," 22 February 2022. [Online]. Available: https://www.oecd.org/en/publications/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en.html.

[14] Berryville Institute of Machine Learning, "An Architectural Risk Analysis of Machine Learning Systems: Toward More Secure Machine Learning," February 2020. [Online]. Available: https://www.garymcgraw.com/wp-content/uploads/2020/02/BIML-ARA.pdf.

[15] U.S. Food and Drug Administration, "Artificial Intelligence-Enabled Medical Devices," 10 July 2025. [Online]. Available: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-enabled-medical-devices. [Accessed 8 August 2025].

[16] U.S. Food and Drug Administration, "Artificial Intelligence in Software as a Medical Device," 25 March 2025. [Online]. Available: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device.

[17] E. M. Bender, T. Gebru, A. McMillan-Major and S. Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?," in FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 2021.

[18] P. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in Proceedings of the 35th Annual Symposium on Foundations of Computer Science (SFCS '94), Santa Fe, NM, 1994.

[19] U.S. Food and Drug Administration, "Cybersecurity in Medical Devices: Quality System Considerations and Content of Premarket Submissions," 27 June 2025. [Online]. Available: https://www.fda.gov/media/119933/download?attachment.

[20] CloudBreach, "Breaking the Chain: How Threat Actors Exploit Supply Chain in Major CSPs," 3 April 2025. [Online]. Available: https://cloudbreach.io/blog/breaking-the-chain-how-threat-actors-exploit-supply-chains-in-major-csps/.

[21] National Institute of Standards and Technology, "Artificial Intelligence - Glossary | CSRC," [Online]. Available: https://csrc.nist.gov/glossary/term/artificial_intelligence. [Accessed 8 August 2025].

[22] International Electrotechnical Commission (IEC), IEC 60601-1 Ed. 3.2. Clause 3.27, 2020.

[23] European Union Agency for Cybersecurity, "ENISA Multilayer Framework for Good Cybersecurity Practices for AI," 7 June 2023. [Online]. Available: https://www.enisa.europa.eu/publications/multilayer-framework-for-good-cybersecurity-practices-for-ai.

[24] U.S. Food and Drug Administration, "Software as a Medical Device (SaMD)," 4 December 2018. [Online]. Available: https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd.

[25] U.S. Food and Drug Administration, "Postmarket Management of Cybersecurity in Medical Devices," 28 December 2016. [Online]. Available: https://www.fda.gov/media/95862/download.

[26] I. Shumailov, Z. Shumaylov, Y. Zhao, Y. Gal, N. Papernot and R. Anderson, "The Curse of Recursion: Training on Generated Data Makes Models Forget," [Online]. Available: https://arxiv.org/abs/2305.17493.

[27] The MITRE Corporation, "ATLAS (Adversarial Threat Landscape for Artificial-Intelligence Systems)," [Online]. Available: https://atlas.mitre.org.

[28] National Institute of Standards and Technology, "Least Privilege - Glossary | CSRC," 3 September 2025. [Online]. Available: https://csrc.nist.gov/glossary/term/least_privilege.