Chapter 9

Interoperability and Information-Sharing Paradigm for IoT-Enabled Healthcare

Brian Desnoyers[†]

MIT Lincoln Laboratory

Kendall Weistroffer[†]

MIT Lincoln Laboratory

Jenna Hallapy[†]

MIT Lincoln Laboratory

Sandeep Pisharody[†]

MIT Lincoln Laboratory

As the line between clinical and personal health devices is blurred with new personal health technologies, there is a need for secure and reliable integration between enterprise Internet of Things (IoT) networks, private cloud networks, and personal connected health devices. This chapter lays out foundational IoT and cloud information-sharing requirements for healthcare, reviews existing and potential approaches to facilitate this integration, and analyzes methodologies for achieving heterogeneous data interoperability between various IoT sensor ecosystems. We describe the information-sharing requirements for a healthcare system infrastructure, and the corresponding security effects on infrastructure-as-a-service (IaaS) and private cloud solutions for data management. The integration of data from personally used IoT sensors, such as smartwatches and fitness trackers, with clinically collected information accessed by medical professionals introduces further security challenges and ethical issues regarding data ownership, efficient data sharing, and privacy. Many of these challenges emerge from traditional medical record access patterns, such as allowing delegation of data access controls during emergency care and the sheer number of personnel accessing medical data for consultations and support, potentially without the full awareness of the patient. The information sharing requirements for a modern healthcare infrastructure, based on IoT endpoints for data collection and cloud computation and storage, include efficient data sharing, access auditing, data filtering and transformation, as well as customizable delegation of data access management responsibilities. We then enumerate various information-sharing approaches to meet these unique demands for IoT and cloud integration in the healthcare field, along with the associated efficiency, availability, security, and ethical consequences of each approach. IoT devices add to data sharing challenges that exist today, starting with inconsistent connectivity interfaces, such as WiFi, Bluetooth, and emerging communication interface technologies, like 5G. The efficiency and adequacy of these approaches will be examined in further detail through the lens of scenarios and dilemmas that may be common in future integrated IoT and cloud healthcare environments. This review includes both existing and potential implementation approaches for IoT and cloud data sharing, providing specific examples of proposed and established systems with their benefits and limitations. Finally, we analyze approaches for achieving interoperability

† Work submitted in this manuscript was done by the authors as independent researchers. The opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by MIT Lincoln Laboratory.

between various ecosystems of IoT sensors to facilitate heterogeneous sharing of relevant IoT-derived health data between patients, healthcare providers, payers, and other authorized parties. The sharing of big data collected from IoT sensors in real-time will require specialized approaches to achieving this data ubiquity beyond requirements necessary for traditional medical records. With existing interoperability challenges between electronic health record (EHR) platforms, IoT sensor data must integrate with public cloud environments to enable improved clinical decision-making and oversight. This analysis will specifically discuss the data ubiquity, availability, and performance consequences for each interoperability approach while enumerating general best practices for integrating, aggregating, and sharing heterogeneous data across multiple IoT ecosystems and cloud environments.

9.1 Introduction

With the advent and proliferation of Internet of Things (IoT)-based health devices around the world, the distinction between clinical and personal devices is becoming increasingly blurred, resulting in unique information-sharing challenges. The global infrastructure of IoT-based health devices consists of a large number of connected legacy medical sensors, IoT-based personal health devices, and software applications that generate vast amounts of medical data that need to be processed, correlated, and analyzed in near real-time. Given the extensive amounts of data, collecting and aggregating the appropriate data from these systems and performing required data processing and computation requires secure and reliable integration of enterprise IoT networks, public and private cloud networks, and personal-connected health devices.

Healthcare is a relevant case study of IoT-cloud network management because it poses several relevant challenges:

1 the existing legacy of medical records and electronic medical records

2 the blurring distinction between medical devices and consumer devices used in healthcare

3 the balance between privacy and decentralized immediate access to data across healthcare providers.

In this chapter, we detail requirements for secure and reliable data management and the integration of IoT-enabled healthcare.

The chapter is presented as follows. Section 9.2 and Section 9.3 discuss two related efforts that inspire requirements for improved interoperability and contribute to the changing paradigm within healthcare: mobile health (mHealth) and

precision medicine. Section 9.4 presents and discusses the challenges to address regarding the ownership of healthcare IoT data, and other integrated medical data sources. Section 9.5 presents the challenges and strategies necessary for data sharing in an IoT-enabled healthcare ecosystem. Finally, Section 9.6 provides a general architecture for interoperability, and then discusses the current state of health record interoperability standards.

9.2 Mobile Health and the Internet of Medical Things

In today's world, mobile technology is ubiquitous. Handheld devices such as smartphones and tablets provide access to information and communications across the world. According to the Pew Research Center, an estimated 94% of adults living in advanced economies own a mobile phone, with the numbers expected to increase within the next few years [357]. This growth in both users and the use of mobile and wireless technologies over the last few years promises a rise in new opportunities for the integration of mobile health technologies. Mobile health, or mHealth, provides users with mobile self-care through the use of consumer apps, devices, and connections that enable users to capture their own health data [246] and receive personal health interventions. Currently, mHealth provides a broad range of services to users, including survey and questionnaire delivery [359], real-time habit recognition and adherence support [63, 138, 276, 374], and pervasive sensor data collection [299, 376]. Although there is no standardized definition of mHealth, we have adopted the definition proposed by the World Health Organization in this chapter:

> mHealth is the use of mobile and wireless technologies to support the achievement of health objectives [257].

In addition to mobile technology, the emergence of affordable, wearable devices has continued to create new opportunities for mHealth. These wearable devices (commonly referred to as "wearables") provide users with a convenient means to monitor and manage personal health and connect to healthcare providers via telehealth (e.g., remote patient monitoring) [263]. Although a vast majority of general-purpose wearables lack specialized health sensors, they have technology components that can provide functionality akin to that of health sensors, such as motion measurement, body tracking, body balance assessment, and pattern recognition [152]. In a broad sense, this ecosystem of connected IoT-based health devices has been termed the "Internet of Medical Things" (IoMT).

Research studies regarding the efficacy of mHealth interventions and outcomes are limited with current evidence showing mixed results [251]. Thus, continued initiatives to conduct systematic studies on the effectiveness of mHealth are essential in determining whether health-related IoT devices are engaging and providing actual value to users, rather than simply collecting data. If mHealth

devices are failing to provide tangible value outside of data collection, adherence may not be sustainable for the general population [99, 263]. Only after overcoming this challenge in future mHealth deployments can these devices provide users with low-cost and real-time mechanisms for the assessment of personal, clinical, and public health through the collection and analysis of movement, imaging, behavior, social, environmental, and physiological data [99].

9.3 Enabling Precision & Personalized Medicine

Evidence-based medicine is the practice of integrating the experiences and knowledge of a healthcare provider with external clinical evidence and patient needs [317, 318]. This integrated evidence comes in many forms with the responsibility for seeking out the best external evidence falling, at least partially, on healthcare providers. The "gold standard" source of external evidence is the randomized controlled trial (RCT) [318, 349], that often evaluates treatment effectiveness on the population scale through clinical epidemiology. Integration of external data is practiced by many healthcare providers today and has shortened the gap for new clinical research to be widely utilized in medical practice [252]. However, a significant challenge remains as healthcare providers must still decide how new studies pertain to their individual patients at the time of care [252]. Eric Topol uses the example of widely prescribed statin drugs for preventing endpoints, such as stroke and heart attack, to illustrate this challenge [4]:

> Instead of identifying the 1 person or 2 people out of every 100 who would benefit, the whole population with the criteria that were tested is deemed treatable with sufficient, incontrovertible statistical proof.

At the time of Topol's writing, common evidence-based practice often involved prescribing statins for large portions of the population, such as elevated calculated LDL cholesterol levels [4]. In the future, this approach could expand with the widespread use of polypills, such as those containing a statin along with aspirin and folic acid [234]. While this approach could be considered better than the alternative of a non-evidence-based practice, providing care based on population risk factors determined from a limited set of data can expose a population to unnecessary side effects (the use of statins has been associated with diabetes mellitus, liver damage, muscle damage, and central nervous system complaints [358]) while adding financial burden to the healthcare system.

A simplified practice of personalized medicine has been commonly deployed through pharmacogenetics to tailor drug prescriptions based on genetic markers [211]. However, it is the integration of IoT and mobile data sources that will allow for the inclusion of behavioral (e.g., activity patterns, habit detection) and environmental (e.g., noise exposure, air quality) data into this process. For example,

Joshi et al. [184] describe the integration of IoMT data sources into the neonatal sepsis prediction process.

Current clinical practice has utilized coarse population models to improve patient care. One such example is antibiograms, which identify local patterns of antibiotic resistance. Clinicians currently apply these localized resistance profiles to best identify the antibiotics to prescribe to patients [208], such as within a hospital setting. A precision medicine approach that improves upon this practice might incorporate additional features, such as social network and location check-ins, and provide these insights for the individual patient. Another example of such a clinical model is the Breast Cancer Risk Assessment Tool (BCRAT), which is a model used clinically to determine breast cancer risk based on factors such as age, race, and family history [268]. This BCRAT risk score can be used to determine recommendations, such as secondary prevention screenings. Precision medicine approaches can yield similar models that can be applied on a more individualized basis, and are an area of current research [268].

The additional real-time data collection and processing enabled by mHealth and IoMT devices will make it increasingly possible to quantify human beings such that the practice of evidence-based medicine can be individualized. Human quantification and data integration will present healthcare providers with additional tools to scientifically determine which patients are the most similar to their own, and thus perform "real-time" epidemiological research to decide on the best treatment using N-of-1 trials [224]. With the aforementioned tools, healthcare providers may be able to identify smaller segments of the population that need particular treatments with higher confidence. When data is integrated on large scales (e.g., the entire population of the United States), over long periods of time, it could potentially become practical to evaluate endpoints of interest (e.g., heart attack) rather than surrogate endpoints (e.g., blood cholesterol levels) to more effectively evaluate treatments. This practice of integrating large sources of genetic (e.g., DNA sequencing), behavioral, and environmental data for developing precise personalized treatment plans is known as precision medicine.

Inclusion of heterogeneous big data sources has the potential to have a transformative effect on evidence-based medical practice and allow for improved healthcare delivery [176]. However, this must be preceded by studies of the efficacy and sustainability of precision medicine interventions [176]. Enabling precision medicine to advance evidence-based medical practice is therefore a significant motivating factor for the deployment of IoT and cloud integration in healthcare.

9.4 Health Data Ownership in IoT and the Cloud

Technological advances within the healthcare industry (e.g., mHealth) have created an unprecedented amount of user-generated, health-related data [199,

257]. Data ownership and the implications on personal and data privacy from third parties attempting to connect to these devices in order to access, capture, analyze, and share this data remains an under-explored area from a policy and regulatory perspective [199, 237]. Thus, there is a clear need for transparent regulations and requirements for health-related data ownership and sharing.

9.4.1 IoT Data Ownership Challenges

Although ownership and protection of health data is an obvious concern within the healthcare field, the use of IoT-based health devices makes the issue more complex. In particular, some of the factors that further complicate establishing data ownership include the unobtrusive nature of the IoT device and its portability, and users' mobility, patterns, and preferences. Further, a vast majority of IoT devices feature unconventional user interfaces, which increases the difficulty of performing a large number of tasks (e.g., user consent and authentication). Given the diverse nature of IoT user interfaces, there currently is no "one size fits all" solution. Developing unique solutions for each IoT device further complicates the endeavor of developing standardized data ownership regulations.

9.4.1.1 Consent for Data Capture

A multitude of IoT-based health devices are continually capturing data from the outside environment. Due to the nature of continuous data collection, the data resulting from this process has the potential to include data sourced from nonconsensual collection, i.e., data collected from individuals without authorization or informed consent, or data collected from individuals unaware of the data collection. In these scenarios, it is difficult to designate the data owner: should the data belong to the owner of the IoT device, or should it belong to the individual whose data is being captured? While traditional IoT devices, such as smart doorbells, are in use, the owner of the device is often responsible for ensuring that any captured data is not violating privacy laws, irrespective of whether or not they are operating the device. Depending on these local laws, the owner might need to obtain authorized or informed consent from all individuals before they are captured by the device.

However, the ability to obtain authorized or informed consent through an IoT device is especially challenging due to its inherent characteristics (e.g., ubiquity, transparency). As an example, placing a physical sign regarding the data collection policies of an IoT device could easily go unnoticed, thereby not constituting authorization or informed consent from the recorded individual [101]. Whether or not an individual or user were to observe a physical sign, notice, or warning, the unique interfaces used in a vast number of IoT devices may prevent the individual from providing authorized consent. For example, audio interfaces used by voice assistants (e.g., Siri, Google Assistant) may not have a visual interface for the user to provide consent in a trivial manner. The ubiquity of IoT devices makes this

process impractical since a user may need to authenticate and provide consent in a location with a large number of devices. Furthermore, multiple requests for consent have the potential to create user fatigue, thereby making user consent invalid and impractical [101].

In a healthcare setting, several of the challenges associated with obtaining authorized and informed consent of data collection from IoT devices can be resolved as part of the registration process for new patients. Similarly, in clinical research, researchers can obtain informed consent for all data collected by any IoT device used within a study as part of the standard informed consent process approved by their Institutional Review Board (IRB). In addition, IoT devices that collect clinically-relevant data outside of healthcare settings should be designed to only collect identifiable data from their consenting subject. Enabling real-time, on-device data processing may help to overcome this challenge. As an example, the Apple Watch provides a Noise app that performs local audio processing to enable users to understand the sound levels in environments that could negatively impact hearing. The Noise app performs local processing without recording audio content and thus does not currently require consent from inadvertently recorded bystanders in areas where it is available [14]. Additional frameworks for obtaining consent from IoT devices where these methods are impractical have been proposed, such as implementing informed consent through gateway devices– either directly from users or indirectly through a centralized registry [101]. For example, The Privacy Coach [76] scans RFID tags within IoT devices to compare the device's specifications to the user's privacy preferences.

9.4.1.2 Verifying Data Ownership: Local Identity Management and Authentication

Many IoT devices used in healthcare are primarily used as sensor devices to collect data. Since the data collected from these devices is being used in an increasing number of clinical decision-making processes, the sensor data can be manipulated in an adversarial manner to change clinical practice. In a potential future clinical situation without a human in the loop, these data integrity issues can lead to significant security gaps, as studied in the field of adversarial machine learning [171]. Even with a human in the loop, healthcare providers will rely on IoT-collected data to make clinical decisions that may be impacted by false data. This data integrity issue requires mechanisms for local user identity management and authentication in healthcare regardless of whether or not the device can provide direct access to clinical data. In practice, these adversaries could be third parties looking to cause harm, or patients looking to alter data for personal gain, such as a patient working to manipulate his data in order to be prescribed a controlled substance.

The distinct interfaces of IoT devices can make local authentication challenging. Even if alternative devices can be used to authenticate IoT devices through proximity, it can create enormous user burden and fatigue with ubiquitous IoT deployment. Developers of IoT devices that collect healthcare-relevant data must

therefore balance the need for authentication with user burden to promote device compliance and limit security risks.

In addition to IoT devices that produce clinically-relevant data, some devices may consume produced data and provide feedback to users. Secure authentication is increasingly vital for these devices because of their potential to leak personally identifiable clinical data. In particular, these devices could reveal personally identifiable information (PII) that is traditionally respected as sensitive both within and outside of healthcare environments.

In order to assist in securing patient information, several IoT device authentication schemes have been proposed. Although these schemes are not widely accepted, they support the non-traditional interfaces of many IoT devices and could be applied to health devices. These authentication mechanisms vary in their overhead and burden to the user, and therefore may need to be considered on a device-by-device basis. These authentication mechanisms may also need to be used in conjunction with other mechanisms (either as an additional or alternative factor) to meet specific device requirements.

A non-comprehensive overview of several authentication mechanisms relevant to healthcare is listed below:

- **Proximity-based Authentication:** Several IoT device authentication approaches rely on the device's proximity to other user-owned devices, often acting as a physical authenticator. Relying on the prevalence of mobile devices such as smartphones, these mechanisms can require varying degrees of interaction with the user. At the simple extreme, proximity-based authentication might involve automatically authenticating devices within a specific range of proximity. On the more stringent extreme, when user attention is deemed necessary, proximity-based authentication mechanisms can require specific user action. Move2Auth [394] is an example of an interactive authentication scheme, wherein it requires users to perform specific hand gestures with a smartphone near the IoT device in order to authenticate. Because of their potential security vulnerabilities, proximity-based authentication schemes must be carefully evaluated before adoption in healthcare environments. Although elliptic curve cryptography (ECC)-based radio-frequency identification (RFID) IoT authentication schemes have been deployed in healthcare environments, all implementations might not have security requirements that are suitable for healthcare deployment [160]. Proximity-based and other physical authenticators may also be used to de-authenticate after proximity or physical contact has ended.

- **Biological Authentication:** Other IoT devices, especially those that already incorporate biometric sensors, may rely on biological authentication. Some biological factors that are used for authentication include fingerprints, face, heartbeat, iris, or voice [166]. Similar to proximity-based authenticators, some biological authentication mechanisms, in combination with other biological or non-biological factors, can also be used to de-authenticate after the biological factor changes or is removed. For example, IoT fitness

devices can use heartbeat or capacitive sensors to detect when the device is removed and thus should be de-authenticated [368]. Because these factors themselves may be considered personal data, storing them directly on devices for authentication is highly discouraged. Instead, the configuration of IoT devices for biological authentication should utilize raw sensor data to train a mathematical model such that the original personal data cannot be reconstructed. Apple's TouchID technology used in select iPhones, iPads, and computers utilizes a similar approach [1].

- **Proxy-Based Authentication:** Proxy-based authentication relies on a secure channel between a specific proxy and the IoT device [86]. In this approach, a proxy (e.g., a clinician) would verify the user's identity and authenticate the device to the user. This proxy access could be granted to the clinician for a specific device to limit the possibility of this privilege being misused. Generally, when proxy-based authentication is required, devices should not de-authenticate with other factors in order to decrease burden on both the users and the proxies (often healthcare providers).
- **Behavioral Authentication:** The less commonly deployed strategy of behavioral authentication relies on non-biological behavior data to identify a user [286]. By relying on behavioral data, the user authentication process can often be implicit, with limited user burden. Although challenging, the development of novel secure behavioral identification techniques would have payoffs that could significantly improve the usability, and, therefore, the adoption of IoT device authentication. Current implementations have been developed that identify habits based on user data such as phone calls and locations, but are currently only suitable as secondary factors within multifactor authentication architectures [179, 328].

9.4.2 Healthcare Data Ownership

Much of the data currently used by healthcare providers is stored within health information technology systems. This data often takes the form of an electronic medical record (EMR), electronic health record (EHR), or personal health record (PHR). While these terms are often used interchangeably, there are some significant differences to these terms with regards to health data ownership. EMRs are traditionally owned by a single office or organization and are mainly a digital version of a traditional paper medical record [15]. EHRs are similar to EMRs in that they are owned by a healthcare organization, yet differ in the context that EHRs are designed to enable sharing with providers across healthcare organizations [15]. In contrast, PHRs defer ownership to the patient who collects and stores information across healthcare systems [4]. While EHRs are in widespread use today in the United States, personally owned records may become more widespread as patients begin sharing data collected by their own IoT devices with healthcare providers. These record systems are examined in further detail in the remainder of this subsection.

9.4.2.1 Electronic Health Record (EHR)

EHRs have achieved widespread adoption within the United States with 84% of hospital [292] and 53.9% of office-based physicians [11] adopting a basic EHR.[1] Adoption of fully functioning systems with additional features is lower. While the ideal EHR would allow for complete data federation across all healthcare providers, this is far from the case in many current isolated deployments. While there is limited work estimating EHR fragmentation, several studies have explored the extent of incomplete health data in EHR systems [70, 240], which can lead to patient harm such as medication errors [61].

The EHR is a healthcare organization-owned system in which patients have limited ownership of their own data. Although regulation, such as the Health Information Technology for Economic and Clinical Health (HITECH) Act in the United States [66], provides incentive for allowing electronic patient access to EHRs, some data remains unavailable. These restrictions can be beneficial, as direct access to test results has been shown to lead to anxiety and increased rates of patient visits [305]. However, allowing direct patient access has also been linked to increased patient engagement and is highly valued among patients [339, 361].

Although electronic health data is not widely used in research, their use in the clinical decision making process is increasing [167], and patients may be willing to share such data for research purposes [192]. With universal interoperability, the EHR can theoretically enable research leading to significant public-health benefits when fully adopted. For example, data trends can be used to

1 increase accuracy in influenza strain predictions for vaccinations,

2 evaluate treatment effectiveness in specific sub-populations, and

3 identify emerging drug resistance.

However, attempts at EHR interoperability face familiar barriers such as missing data [203]. In scenarios where complete interoperability has not been achieved, the EHR remains a healthcare organization-owned record. This can lead to a lack of efficiency but could also have harmful patient effects, such as redundant imaging [210], as patients visit healthcare providers using different EHR systems. Finally, there is prevailing belief that patients should have a right to accept the consequences to access and manage their own data [361], regardless of the potential risks [305]. This belief is related to the ethical principle of patient autonomy, and accepted practice of informed consent.

9.4.2.2 Personal Health Record

The PHR is a patient-centered form of medical record in which data is stored on a patient-owned portable device, or cloud service that the patient is able to

[1] A basic EHR, as defined by DesRoches et al. [110], includes support for patient demographics, patient problem lists, medication lists, clinical notes, prescription order entry, viewing laboratory results, and viewing imaging results.

access. This ensures that a patient's personal medical records are always available to them. When the patient moves from their primary care provider, to an urgent care clinic, to a medical testing facility, they are able to bring their record with them to be populated regardless of the record system used by the practice.

The PHR can also solve many of the data federation concerns associated with EHRs, as the patient is responsible for maintaining a single record which can be accessed by all care providers. Two well-known PHR systems were Google Health [232, 233] (closed in 2011) and Microsoft HealthVault [10] (closed in 2019), both of which originated from the Personal Internetworked Notary and Guardian (PING) or Indivo systems [12, 247, 332]. These PHR systems provided users with a single portal through which they could access their health information by linking with existing EHRs via interoperability standards such as the Continuity of Care Record [132]. Beyond that, these systems provided functionality for fine-grained data sharing as well as integration with personal health devices and health apps [10]. Unfortunately, the universal PHR systems have waned in favor of fragmentation, reminiscent of the current state of EHR [52]. In this fragmented model, users are connected to healthcare provider-owned EHRs as well as siloed commercial PHR systems such as the ones provided by pharmacies or fitness trackers. This fragmentation has occurred in context of a trend toward patient-driven self-care, including quantified self-tracking (e.g., PatientsLikeMe, 23andMe, Fitbit) [350], resulting in many, often data type-specific, PHR systems. Beyond the inconvenience of maintaining several record systems, the integration of these systems could limit patient harm by decreasing the prevalence of incomplete medical records [70].

However, the tide may be turning back toward a universal PHR system as tools such as Apple Health continue to gain in popularity and achieve high user satisfaction [7, 103, 283]. Apple Health is significantly different from Google Health and Microsoft HealthVault, in that it is a product of the smartphone era with the interface and data localized to an owner's device rather than an Internet portal. The obvious consequence of this is privacy: by retaining health data encrypted locally on a user's device, a user does not need to be as concerned about data misuse or the compromise as with an Internet-hosted portal. However, by hosting the PHR locally on an owner's device, Apple Health is also less able to provide two-way data transport: the app is able to collect data from healthcare providers, rather than currently providing data to providers. While the two-way data transport does not solve the problem of incomplete EHRs, Apple's CareKit [2], a developer framework for applications that allow patients to share health data with healthcare providers is touted as the potential solution. Apple Health utilizes an improved interoperability standard for accessing EHR data from participating healthcare providers [62], which is discussed in detail in Section 9.6.

9.4.2.3 Bridging Medical Data Ownership: Combining EHR and PHR

Currently, healthcare providers continue to maintain and collect patient data within EHRs, and patients collect personal data within (perhaps fragmented)

PHRs. It therefore follows that EHR-PHR interoperability will be an important step for mHealth and personal IoMT data utilization in healthcare. A simplified model of EHR-PHR interoperability that federates data sharing between both healthcare provider and personally owned IoMT devices is shown in Figure 9.1.

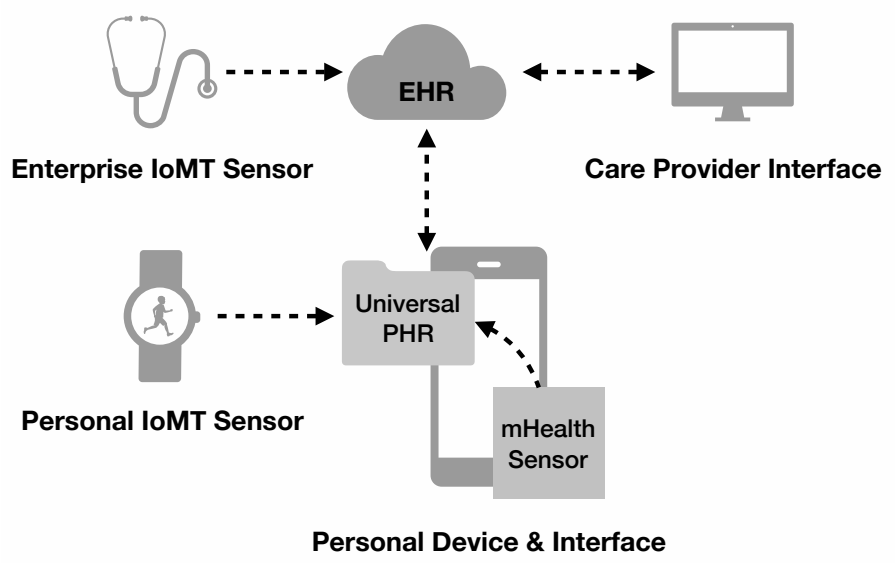

Figure 9.1: A simplified model of interoperability between an EHR and PHR system that can be used to integrate personal and enterprise IoMT devices. The PHR hosted locally on a user's smart phone can be replaced with a cloud-hosted PHR without modifying the model.

While this simplified model directly shows only a single EHR and universal PHR, additional EHRs could connect with the universal PHR via the same mechanism. By linking multiple EHR systems through a single universal patient-owned PHR, interoperability of both healthcare provider-owned and personally owned health records can, in effect, be achieved to prevent gaps in medical data. Additional architecture details for enabling health information exchange, such as cloud-hosted PHRs and 5G IoMT sensors, are discussed in Section 9.5.

In addition to mitigating EHR silos, combining EHR and PHR systems in this fashion also addresses a significant limitation associated with PHRs. As discussed previously, EHRs provide repositories of aggregated patient data that can be mined for public health and precision medicine research. Unless PHRs are hosted together, the PHR may lack the ability to serve as a research data repository [369]. While many EHR systems share only basic amounts of data, some large EHR networks, such as Epic [6], maintain interoperability between their EHR systems by including medical data that may be derived from IoMT devices. The integration of PHRs with these large connected EHR systems can produce data that would allow for public health research. It should be noted that

access to care from providers that are part of these EHR network-based data warehouses may be limited for some populations, such as those with unequal access to care. Until universal access to these systems is established, clinical use of this data must therefore account for this skew.

9.5 Enabling IoMT Information Sharing in Healthcare

This section examines the strategies necessary for the federation of IoMT data in healthcare. The section starts by focusing on strategies for the health information exchange of traditional medical records. The section then discusses several challenges with federating IoMT data via similar mechanisms as well as current solutions, where available.

9.5.1 Collecting Data from IoMT Devices

Logically, the first step in being able to use data from IoMT devices is to collect the data generated by the IoMT devices and sensors. Since the IoMT devices used for support operations in hospitals are entirely enterprise owned, we will focus on IoMT devices operating within their 'clinical' use case.

At a high level, the collection of healthcare data from IoMT devices, whether the devices are enterprise owned or personally owned, has broad similarities. In both cases, there are two main considerations that need to be addressed for effective data collection. First is managing the enormous amounts of data that IoT devices are capable of generating. Decisions need to be made on what data is collected and the frequency of data collection. Several aspects of personalized medicine depend on integrating vast amounts of collected medical data for clinical research. However, this need should be balanced with identifying the important pieces of data since the amount of data collected also directly ties into the network resources required to transmit the data to the cloud or an alternate data storage location. Data thinning techniques should be applied to retain only essential data, so as to help reduce overhead in transport and data processing needs at a later stage. Second, and more important, is security. Personal data transmitted by devices that monitor health must be secure to protect personal privacy.

However, the primary differentiation between enterprise IoT devices as opposed to personal healthcare devices and Mobile IoT sensors is the user. The average user is not knowledgeable enough to make decisions on the magnitude of determining the collection frequency for personally owned IoT devices. Users also tend to be more at-risk for security exploits than enterprises practicing good software hygiene.

9.5.2 Traditional Health Record Information Exchange for Information Federation

Health information exchange is the sharing of patient-level electronic health information for assessment, cost reduction, and quality improvement in healthcare [369]. The HITECH Act mandates a limited level of health information exchange to be eligible for incentive payments in the United States [369]. Traditional information formats for medical record information exchange are discussed in Section 9.6.2. While healthcare provider relationships can enable exchange without significant technological access control requirements, additional considerations must be made when sharing data between patients and healthcare providers. The methodologies for traditional medical record information exchange can vary across emergency and non-emergency situations, and provide insights for the integration of personal IoMT and mHealth data into EHR systems [344].

9.5.2.1 Regulating Provider Access to PHR Data

Multiple fine-grained access control methods for PHR data have been described to provide secure information sharing of health data. Such an approach would allow for patients to have control over which users are able to access specific information contained within the patient's encrypted PHR. In order to enforce such control over their health records, patients would have the authority to generate and provide decryption keys based on the information they wish to provide to the receiving party. This method of access control would enable the secure sharing of PHR data with authorized healthcare providers while protecting the patient's personal data from unauthorized parties. However, many of these fine-grained access control methods result in high overhead costs when applied to scenarios involving multiple users, and thus impacting system usability. A fine-grained access control framework for PHR data with reduced overhead has been proposed by Li et al. [220]. This approach involves users generating their own sets of attribute-based encryption (ABE) keys. To account for the linearity of ABE encryption, the system is divided into multiple domains which are associated with various user subsets.

9.5.2.2 Providing Emergency Data Access

Although access control of medical records can be achieved through the previously discussed methods, in the case of required emergency access to data contained within the PHR and EHR, personal health data may become available without prior authorization. Currently, emergency data access to protected records within a single EMR system often follows a "break the glass" procedure. This procedure involves a healthcare provider self-granting access to a patient's medical record and protected health information (PHI) that can be utilized in the event of an emergency. Each instantiation of the "break the glass" procedure is documented and is later audited and reviewed to ensure that the patient's

medical record and PHI were accessed under justifiable circumstances. Because an auditing, review, and accountability process exists under the "break the glass" procedure, it is not clear how this approach can be implemented for the sharing of health records across organizations using separate EMR systems or for granting access to PHRs in the case of emergencies.

The current practice of utilizing the "break the glass" procedure in emergency circumstances presents the risk for providing unnecessary access to patient information. One method of differentiating between users who should be granted emergency access to patient data and users who should be denied access integrates both Role Based Access Control (RBAC) and Experience Based Access Management (EBAM) strategies [398]. In order to test the effectiveness of this approach, the resulting algorithm, "Roll-Up", was applied to log data collected from Northwestern Memorial Hospital Center. Results from this case study indicate that a combination of RBAC and EBAM strategies is able to predict the conceptual position of a user requesting access to a patient's EMR data with 82.3% accuracy [398].

In the case of Li et al.'s proposed fine-grained access control framework, the issue of handling the security risks associated with providing data access during an emergency is handled using decryption keys [220]. This trapdoor method involves the patient selecting which parts of their PHR data they wish to be accessible in advance of a health emergency. The patient is able to delegate access of this data to the emergency department by providing a decryption key for each part of the pre-selected PHR data. These decryption keys would be stored within the emergency department's database of patient information. If an emergency occurs, a staff member would be able to query the database and obtain the patient's decryption keys from the emergency department. Once the patient's medical condition has returned to normal, the patient's PHR system could then compute re-keys for their PHR data and submit this update to the emergency department for future use. Although naturally supported by the framework proposed by Li et al., it remains unclear if this "break the glass" method would be able to scale and work across multiple hospital locations, given that the patient must be able to provide decryption keys to a particular emergency department in advance of any emergency incident. In order for this emergency data access method to be used across hospitals, the decryption keys provided by the patient would have to be stored in a centralized database, causing a host of other security and scalability issues.

Digital Rights Management (DRM) schemes can also be considered as a way to secure PHR and EHR data from unauthorized insider access. Kunzi et al. propose a data-centric model for the protection of health records in which encrypted health data is able to be accessed in an emergency with use of an emergency license [194]. Under such circumstances, an emergency key is issued in order to decrypt the patient's health data. Similar to the "breaking the glass" protocol, emergency access is documented and audited to ensure appropriate record access. However, in order to prevent against system compromise, a compromised emergency key will have a limited effect on the system. Additionally, the

system design of this approach also ensures the dependability of the system while operating offline.

9.5.3 Ensuring Data Integrity from IoMT Sensors

Unlike traditional EHR-integrated sensors, the collection of data from mHealth and IoMT devices involves connections to large numbers of devices outside of the control of the healthcare organization. These devices may integrate via local gateways, connect directly to the EHR, or connect to cloud-hosted data warehouses. Regardless of the connection mechanism, the data collected from these devices will integrate with a health record system (likely an EHR or PHR). With potentially numerous connected devices, it is important to maintain a device inventory and validate that unaltered sensor data is being transmitted and received.

A simple mechanism to mitigate integrity risks for direct sensor device transmissions would be to utilize cryptographic encryption and signing, with validation performed based on public keys stored within a centralized device inventory. In this model, an mHealth or IoMT device would periodically be registered with the health record system, in which the device generates a key pair and shares the public key with the system's sensor device inventory. The mHealth or IoMT sensor would also receive a public key to encrypt the data sent to the health record system. The sensor could then encrypt and sign traffic to the health record system, which could be decrypted and then validated. The health record system would then tag the data source for each piece of received data. This method would help to ensure only data from valid, inventoried sensors is shared with the health record system, and provide a basic method for tracking the sources of received sensor data.

9.5.4 Privately Replicating and Sharing Large Datasets

In the EHR-PHR interoperability mechanism discussed previously, data is replicated and stored across numerous health record systems. As collected data sources become larger, such as when dealing with genome sequences, replicating data in its entirety across several cloud systems becomes impractical. A more efficient solution would be for record systems to use a "link" to a single instance of the data. The record systems which do not store the data in its entirety could initially download the data and compute any summary statistics necessary to store locally, before deleting the data. These calculated summary statistics might be used directly by the organization, such as common single nucleotide polymorphisms (or SNPs) from genomic data, or the number of steps traveled from motion data. Future calculations or analysis could be performed by simply downloading the file from the "link" again, or utilizing an application programming interface (API) provided by the file host.

This solution for enabling the sharing of large datasets does not address the ownership of the large file that is linked to by other sources. Naively, it could be proposed that the owner of the connected device that provides the data must also

host these files. For example, if a patient received a genetic test from 23andMe, then 23andMe would be responsible for hosting the results indefinitely. While it is clear how this would work with tests from 23andMe or imaging from a healthcare provider, certain other scenarios aren't quite so straightforward. For example, how would motion data from personally owned devices be stored? What would happen if the company hosting the data went out of business? Would a user be responsible for paying to host their own data in order to receive the best care? Furthermore, when a single entity is responsible for hosting a data file, ensuring redundancy and availability can be prohibitively expensive.

An alternate solution is for these data files to be hosted using a consensus mechanism, in which small pieces of data are stored in a distributed fashion across several source hosts. These data pieces can be stored redundantly and in fault-tolerant fashion, such that if a single data host becomes unavailable, its shards would remain available from other hosts [59]. As a single file is replicated across additional sources, these sources would individually need to store less of the overall data. These data hosts can include EHR providers, commercial data providers or device manufacturers (e.g., 23andMe or Fitbit), and user-owned cloud storage that may be reimbursed by insurance providers or governments. Some published mechanisms for achieving this distributed data storage method would be suitable for this application in healthcare. The Security-Aware Efficient Distributed Storage (SA-EDS) model proposed by Li et al. requires data packets to be retrieved from a set of cloud storage providers before yielding the original data [219]. Similarly, Shafagh et al. propose a blockchain-based mechanism for secure distributed storage and sharing of time series data [325] that can also be modified to include genomic or other large non-time series data.

9.5.5 Maintaining Consensus in Large-Scale Federated Systems

Not all IoMT data is large enough to require efficient distributed storage. For smaller data sets that can be replicated across several medical records, there is an opportunity for maintaining consensus. Today, much of this burden falls on the patient. For example, although vaccination records are shared between healthcare providers, it is likely to be the patient who would catch a healthcare provider mistakenly administering a vaccination that they have already received. Often, these small mistakes may go unchecked, but, when caught, may lead to changes in treatment. For example, in a 2004 study of a computerized medication reconciliation tool, physicians changed the discharge orders for 94% of patients when discrepancies were identified by nursing staff [308].

Traditional software-only consensus solutions are inadequate in such situations as patients may report or present different information to different providers. For example, a patient might not take a medication prescribed by one physician and might not report (intentionally or unintentionally) taking the medication to another. In situations such as these, having the medical records retain a history of change would be a welcome feature. This would allow medication reconciliation software to automatically detect the discrepancy and allow

healthcare providers to clarify information with the patient to update the record, which will retain the complete history of the reconciliation.

In medical records, retaining complete history is a necessary step for medical record reliability as allowing for the free deletion of medical records can allow patients to significantly affect the behavior of healthcare providers [381]. To ensure record reliability, changes to medical record values should not be allowed in EHR systems. In patient-controlled PHR systems, updated records should be assigned a new identifier such that they are shared with EHR systems as a new value. While some laws may locally require that patients are able to delete medical data stored in EHR systems [381], these deletions do not need to propagate– thus requiring patients to request their records be deleted across all EHR providers. This limitation would help prevent against the compromise of medical records within a large-scale federated health record system.

9.5.6 Providing Emergency Access to Real-Time IoMT Data

The collection and access of real-time IoMT data currently presents a potential challenge in the event of an emergency. IoMT devices are able to collect a variety of health data from users, including biometric data such as a user's heart rate, physical activity, and sleep cycle. Due to the high sampling rates of IoMT health monitoring services and applications, personal health monitoring devices have the ability to produce large sets of health data for each user. In the case that emergency access to this data is required for medical treatment, the volume of available patient data could result in finding the discernible 'signal' within this big data noise to be a challenging problem requiring the retrieval, sorting, and selection of relevant data.

Among these challenges is the issue of ensuring that the real-time data retrieved by the emergency department is the most up-to-date data collected from the patient's IoMT device. Currently, these devices may not sync to provide updates to a patient's EHR on a timely basis. Thus, the IoMT data retrieved under emergency circumstances may not be an accurate representation of the patient's current (or near-current) state of health. In order to obtain access to the patient's real-time IoMT data, the Emergency Department would likely have to be able to physically access the patient's personal IoMT collection device. This alternative access method may not be feasible in emergency situations as it relies on the assumption that a patient is conscious and able to give consent to allow for the emergency department to access to their IoMT device. Even if consent for emergency access to a personal device is granted, such access poses the risk of creating additional patient privacy concerns.

A potential path forward is for IoMT devices to enhance data availability by having the user select an interval where the device would update the user's real-time health data and offer a selected subset to emergency departments. This data update interval could include a range of data transfer periods for user selection. For example, when initially setting up their IoMT device, a user could decide to have their IoMT device automatically send their collected data to a selected

emergency department(s) in hourly, daily, or weekly intervals. The emergency departments could then choose the interval in which previously received, and now out-of-date, data is discarded.

9.6 Achieving Heterogeneous Data Interoperability

Data interoperability in healthcare is a challenging problem even without the added complication of incorporating IoMT devices as data sources. While this interoperability is simpler in some countries which utilize nationally standardized EHR systems, interoperability challenges can still exist between jurisdictions. Outside of these areas, EHR fragmentation is exacerbated by enterprise-hosted deployments of proprietary and custom software that makes interoperability a significant challenge, as discussed earlier in this chapter. While interoperability standards do exist, these standards are often a subset of the data stored within a single EHR system. The integration of personal IoMT devices into this system only adds to the problem, requiring more comprehensive, flexible, and centralized integration standards.

This section will provide a general architecture for interoperability, and then discuss this current state of health record interoperability standards. Finally, it will describe potential alternative approaches for the interoperability of health data that can more directly enable IoT-cloud information sharing in healthcare.

9.6.1 Interoperability Architecture Overview

Compilers used in software development have a challenging interoperability task– to convert human-understandable source code to executable machine code that can be run on multiple target machines. At a high level, this process is broken into a front-end and a back-end, with the front-end yielding an intermediate representation of the source code, and the back-end converting that intermediate representation to machine code. It is this intermediate representation that prevents the need for a 1:1 match of compilers for every combination of languages and target computers, allowing the front-end to focus on the programming language and the back-end to make optimizations for the target machine.

Similarly, it would be problematic if promoting interoperability in healthcare required a specific tool to provide interoperability between each set of systems. The sheer magnitude of this task would make it impossible to enforce widespread interoperability via policy or encourage interoperability through market forces. Thus, IoT-cloud information sharing in healthcare would require a set of intermediate standards for interoperability. Such interoperability standards would need to be flexible enough to support the variety of data that can be made available through IoMT and mHealth devices.

While there are several more comprehensive models of interoperability, two commonly discussed variants of interoperability include syntactic interoperability and semantic interoperability [81]. Syntactic interoperability refers to the format and encoding used when data is transferred [81], such as the eXtensible Markup Language (XML) format used by Health Level 7's Clinical Document Architecture (CDA) [62]. The syntactic interoperability format would not help interpret the data, but provides the basic foundation structure necessary for interoperability. A set of modern, commonly accepted formats for syntactic interoperability, such as XML and JavaScript Object Notation (JSON), are already being used in similar applications and have additional widespread use outside of healthcare.

Semantic interoperability solidifies the meaning of data such that the meaning is not ambiguous as it is transferred, either between systems or humans [81]. This can be a challenge as health information can be represented both through different names or via a different structure in different systems. For example, one EMR system may represent a heart rate reading as a diagnostic event, while a PHR system might include heart rate information in a different structure designed for fitness data. Interoperability between these systems could be achieved by ensuring that they could both read and write to a comprehensive and flexible intermediate file format with its own representation and designation for heart rate data [356].

9.6.2 Current Interoperability Standards

Two commonly used examples of interoperability standards which dictate the formatting and exchange of health data are The American Society for Testing and Materials' Continuity of Care Record (CCR) and Health Level 7's Clinical Document Architecture (CDA). The purpose of both the CCR and CDA was to support healthcare data management and transfer between healthcare providers. Although the CCR and CDA were both created with the purpose of facilitating the collection and transfer of medical documentation between providers, and are specified in XML, they differ enough that they continue to co-exist.

The CCR was created specifically as a way for healthcare providers to collect patient information in an organized format that can easily be transferred between multiple care providers [132]. The resulting set of standards focuses on the patient's current state of health and other information, such as health insurance, care documentation, and practitioners. A key aspect of the CCR is its focus of presenting patient information in an easily human-readable format. A CCR document is divided into six sections including: an XML header, patient identifying information, patient financial and insurance information, the patient's health status, care documentation, and care recommendations. Although uncommon, IoMT providers are able to allow for the use of CCR data with their devices. Prior to the discontinuation of Google Health and Microsoft's HealthVault in 2011 and 2019, respectively, users of these services were able to integrate both personal health devices, health apps, and CCR data [3, 16].

As noted, the CDA was created to serve a similar purpose: to provide a method of standardization for the storage and transfer of healthcare documents. However, the CDA differs from the CCR on the aspect that the CDA was created to include different levels of machine readability in order to facilitate the transfer of documents between devices. In total, three distinct CDA levels exist: Level 1 is considered to be the most suitable formatting for older systems as it allows for an unstructured free text to be transferred between systems, Level 2 adds structure to the transferred documents by requiring the body of the document to be specified in XML, and Level 3 allows for the highest level of machine readability by requiring an encoded XML document [132]. Simply put, the interoperability of CDA specified documents increases with each subsequent CDA level.

9.6.3 Future Standards and Alternative Methods

Today, there are an increasing number of health-related mobile applications and IoMT devices being produced and used by consumers. Although these devices encourage users to take responsibility for the storage and use of their medical data, these personal devices may require both their users and healthcare providers to access brand-specific applications in order to retrieve health data. In addition to personal IoMT and mHealth devices, prescribed connected medical devices (e.g., remote cardiac monitoring portfolio from Abbott) also utilize web-based portals for healthcare provider access [9]. While some manufacturers provide health record integration solutions, such as Abbott's EHRDirect export [9], these integrations might not be feasible for some healthcare providers. This use of custom interfaces creates additional work for clinicians who wish to access patient data for use in treatment processes, serves as a barrier to data analysis tools, and emphasizes the growing need for updating the current interoperability standards.

To account for the rapid advances of the mHealth and IoMT device industry, interoperability standards need to be revised to allow for a greater degree of flexibility. As discussed in previous sections, the current interoperability standards do not allow for the easy addition of new healthcare data collected from personal devices. Future alterations to current standards could allow for mHealth and IoMT devices to automatically add recently collected data to a patient's medical record. This approach could introduce a standard way for both current and new device companies to provide updated user data to healthcare professionals without the use of custom applications for data access and retrieval. In addition, allowing for the device to update a patient's record automatically would remove the responsibility of providing a personal device's collected data to emergency departments from the patient. Such revisions to current interoperability standards could include methodologies that focus on incorporating a flexible editing and update process, or could take a more iterative approach.

Altering current interoperability standards to permit editing and updating of processes could allow for data to be added from mHealth and IoMT devices without the use of explicit standard updates being released. In this case, a newly

released device would be able to add the user's data to their health record in the form of a device-specific field, which could then be updated with the user's most recent data in the future. These fields would ideally follow a standard format and could include information about the device itself in addition to the user's data such as: device manufacturer, device name, and date and time of last data update. Previously, larger companies such as Google and Microsoft have been able to provide integration of their devices with commonly used health record standards [3, 16]. An increase in flexibility which allows a way for mHealth and IoMT devices to add to and update a user's EHR in a standardized way has the potential to provide the same level of data sharing support to health device companies regardless of reputation, popularity, and size.

Similarly, the development of a universal, accepted set of standards would enable additional semantic interoperability of data for new IoMT and mHealth devices, and enable integration with EHRs. Lopez and Blobel have defined a development framework that could serve as a starting point for the design of such a standard relying on the Rational Unified Process (RUP) framework [235]. As such a semantic standard would be required to be comprehensive; an iterative framework similar to RUP may be able to allow for expansion during the development process. Despite their intended comprehensiveness, semantic interoperability models developed during these processes should be flexible to allow for the integration of new data types that may be similar to data already designed to be incorporated into the model. An example of such flexibility might be allowing a new fitness measurement recorded by a novel IoMT device to be recorded along with other fitness measurements in the interoperability format, along with a name and description of the previously unknown measurement type. Eventually, if this device and new fitness measurement become popular, this data can be added to future iterations of the standardized format. Reconciling data stored in this way with expanded standardized formats can be achieved using thesauri or word embedding, similar to the MetaNet technique developed for metadata domains [173].

9.7 Challenges & Opportunities

Precision medicine, as made possible by the advent and proliferation of IoMT devices and widely available genomic data sharing, holds great promise. However, success in being able to provide individualized evidence-based medical care is dependent on several key issues being addressed. Primary amongst these are:

1. Secure and reliable integration of data from clinical and personal health devices between enterprise IoT networks, private cloud networks, and personal connected health devices.

2. Since research regarding efficacy of mHealth interventions and outcomes is limited, continued initiatives to study the effectiveness of mHealth is essential to determining whether health-related IoT devices are engaging and providing actual value to users, rather than simply collecting data.

3. Settling ownership of data generated from IoMT devices is paramount to getting buy-in from customers wary of privacy violations.

4. A wide variety of IoMT devices makes achieving interoperability between these sensors to facilitate heterogeneous sharing of relevant IoT-derived health data between patients, healthcare providers, payers, and other authorized parties absolutely essential.