A SMART AND INTERACTIVE EDGE-CLOUD BIG DATA SYSTEM

by

Jake Stauffer

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Electrical and Computer Engineering at IUPUI

Indianapolis, Indiana

August 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Qingxue Zhang Chair
Department of Engineering and Technology

Dr. Brian King
Department of Engineering and Technology

Dr. Shiaofen Fang
Department of Computer and Information Science

Approved by:
Dr. Brian King

Dedicated to my beautiful wife Kaia, my best friend Kevin, and my family for their love, support, and encouragement.

ACKNOWLEDGMENTS

Having gone through my graduate year of study, I have learned a lot and spent my graduate year gaining understanding and writing this thesis on an interactive edge-cloud big data system. I am very thankful for my thesis adviser, Dr. Qingxue Zhang and his help with my research. I would also like to thank Dr. Brian King and Dr. Shiaofen Fang for serving on my thesis committee. Likewise, I want to thank Sherrie Tucker who was a tremendous help to me as graduate student and even as an undergraduate at IUPUI. Last but not least, I am incredibly thankful for the support of my family. Their unconditional love, encouragement and sacrifices helped me get through all times of my graduate year. I could not have achieved this without support from them.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Data and information have increased exponentially in recent years. The promising era of big data is advancing many new practices. One of the emerging big data applications is healthcare. Large quantities of data with varying complexities have been leading to a great need in smart and secure big data systems.

Mobile edge, more specifically the smart phone, is a natural source of big data and is ubiquitous in our daily lives. Smartphones offer a variety of sensors, which make them a very valuable source of data that can be used for analysis. Since this data is coming directly from personal phones, that means the generated data is sensitive and must be handled in a smart and secure way. In addition to generating data, it is also important to interact with the big data. Therefore, it is critical to create edge systems that enable users to access their data and ensure that these applications are smart and secure. As the first major contribution of this thesis, we have implemented a mobile edge system, called s^{2}Edge. This edge system leverages Amazon Web Service (AWS) security features and is backed by an AWS cloud system. The implemented mobile application securely logs in, signs up, and signs out users, as well as connects users to the vast amounts of data they generate. With a high interactive capability, the system allows users (like patients) to retrieve and view their data and records, as well as communicate with the cloud users (like physicians). The resulting mobile edge system is promising and is expected to demonstrate the potential of smart and secure big data interaction.

The smart and secure transmission and management of the big data on the cloud is essential for healthcare big data, including both patient information and patient measurements. The second major contribution of this thesis is to demonstrate a novel big data cloud system, s^{2}Cloud, which can help enhance healthcare systems to better monitor patients and give doctors critical insights into their patients' health. s^{2}Cloud achieves big data security through secure sign up and log in for the doctors, as well as data transmission protection. The system allows the doctors to manage both patients and their records effectively. The doctors can add and edit the patient and record information through the interactive website. Furthermore, the system supports both real-time and historical modes for big data management. Therefore, the patient measurement information can, not only be visualized and demonstrated in real-time, but also be retrieved for further analysis. The smart website also allows doctors and patients to interact with each other effectively through instantaneous chat. Overall, the proposed s^{2}Cloud system, empowered by smart secure design

innovations, has demonstrated the feasibility and potential for healthcare big data applications. This study will further broadly benefit and advance other smart home and world big data applications.

1. INTRODUCTION

The age of data is continuing to press forward. Many aspects of our society are being consumed by and are generating data. Additionally, our society is ever more dependent on this data to make decisions. The sheer amount of data that is being produced and consumed has led to what is known as big data. Big data consists of many unique characteristics like volume, velocity, and variety.

With our reliance on data, new challenges and problems are arising due to this increasing amount of complex big data. These challenges have sparked a great need for development in smart and secure data systems. Therefore, it is necessary to develop systems that can handle the big data characteristics listed above. Technologies to manage this big data have been developed and are continually improving.

We take special interest in the healthcare big data [1-4], and propose innovative big data systems correspondingly in this thesis. The research can also be applied to other big data areas. To advance the healthcare big data applications, both edge and cloud systems are crucial. The edge system is critical to capture the big data and the cloud system is important for doctors to manage the patient information and the big data. It is worth noting that traditional electronic health record systems usually do not stream the big data.

The edge system [5-8], especially the mobile phone, has been ubiquitous nowadays. Ever since their development, smartphone usage has grown across the globe. These devices are always with us and can provide unique insights into our lives. These insights can be found within the data that is captured on the phone from the embedded sensors in them. This makes smartphones a natural source of big data that can be very useful in our data driven culture. The smart and secure mobile edge system is obviously important to advance the healthcare big data applications. Therefore, the first major contribution of this thesis is to design and develop an innovative mobile edge system, which will be detailed in Chapter 2.

Another major contribution of this thesis is to design and develop a cloud big data system [9-12]. It is not only essential for the doctors to access and manage the patient data, but also important for enabling effective doctor-patient communications. The details will be given in Chapter 3, which gives the detailed introduction, proposed system, big data security design, interactive chat, patient management, record management, big data management, big data visualization, as well as details results and summary. The developed system is innovative, which allows for secure management of both doctor and patient information, and supports both real-time and historical visualization of the big data.

The overall aim of this thesis is to demonstrate the cloud & edge-based, real-time, smart and secure system that will enable the effective big data management and cloud-edge communications. This thesis, after detailing the mobile edge system and the cloud big data system, will then conclude the research findings and give future studies.

The structure of the thesis is as follows:

Chapter 1. Introduction

Chapter 2. Mobile Edge for Smart and Secure Big Data Interaction

Chapter 3. A Novel Cloud System for Smart and Secure Big Data

Chapter 4. Conclusion

Chapter 5. Future Studies

2. MOBILE EDGE FOR SMART AND SECURE BIG DATA INTERACTION

Data and information have increased exponentially in recent years. Large quantities of data with varying complexities have been leading to a great need in smart and secure big data systems. Mobile edge, more specifically the smart phone, is a natural source of big data and is ubiquitous in our daily lives. Smartphones offer a variety of sensors, which make them a very valuable source of data that can be used for analysis. Since this data is coming directly from personal phones, that means the generated data is sensitive and must be handled in a smart and secure way. In addition to generating data, it is also important to interact with the big data. Therefore, it is critical to create edge systems that enable users to access their data and ensure that these applications are smart and secure. In this study, we have implemented a mobile edge system, called s^2Edge. This edge system leverages Amazon Web Services (AWS) security features and is backed by an AWS cloud system. The implemented mobile application securely logs in, signs up, and signs out users, as well as connects users to the vast amounts of data they generate. With a high interactive capability, the system allows users (like patients) to retrieve and view their data and records, as well as communicate with the cloud users (like physicians). The resulting mobile edge system is promising and is expected to demonstrate the potential of smart and secure big data interaction.

2.1 Mobile Edge for Smart and Secure Big Data Interaction

The amount of data that is generated by a single person has grown exponentially and smartphones are an important contributor to the increase in amount of data. This vast amount of complex data is known as “big data”. These phones have a wide variety of sensors available to them and are therefore enabling easy access to raw data that is tracking the user [13, 14]. This gives rise to an opportunity in many fields, and one typical application is mobile health (mHealth). The role of mobile and remote health monitoring is becoming more and more popular [15, 16], and the demand of ubiquitous healthcare has ever increased. Patient health is very important and technology has enabled doctors to more successfully monitor and improve the health of their patients. A major way to monitor and improve someone’s health is to make informed decisions on the data that is collected. An easy source of this data are devices that many patients have and use all day – namely smartphones.

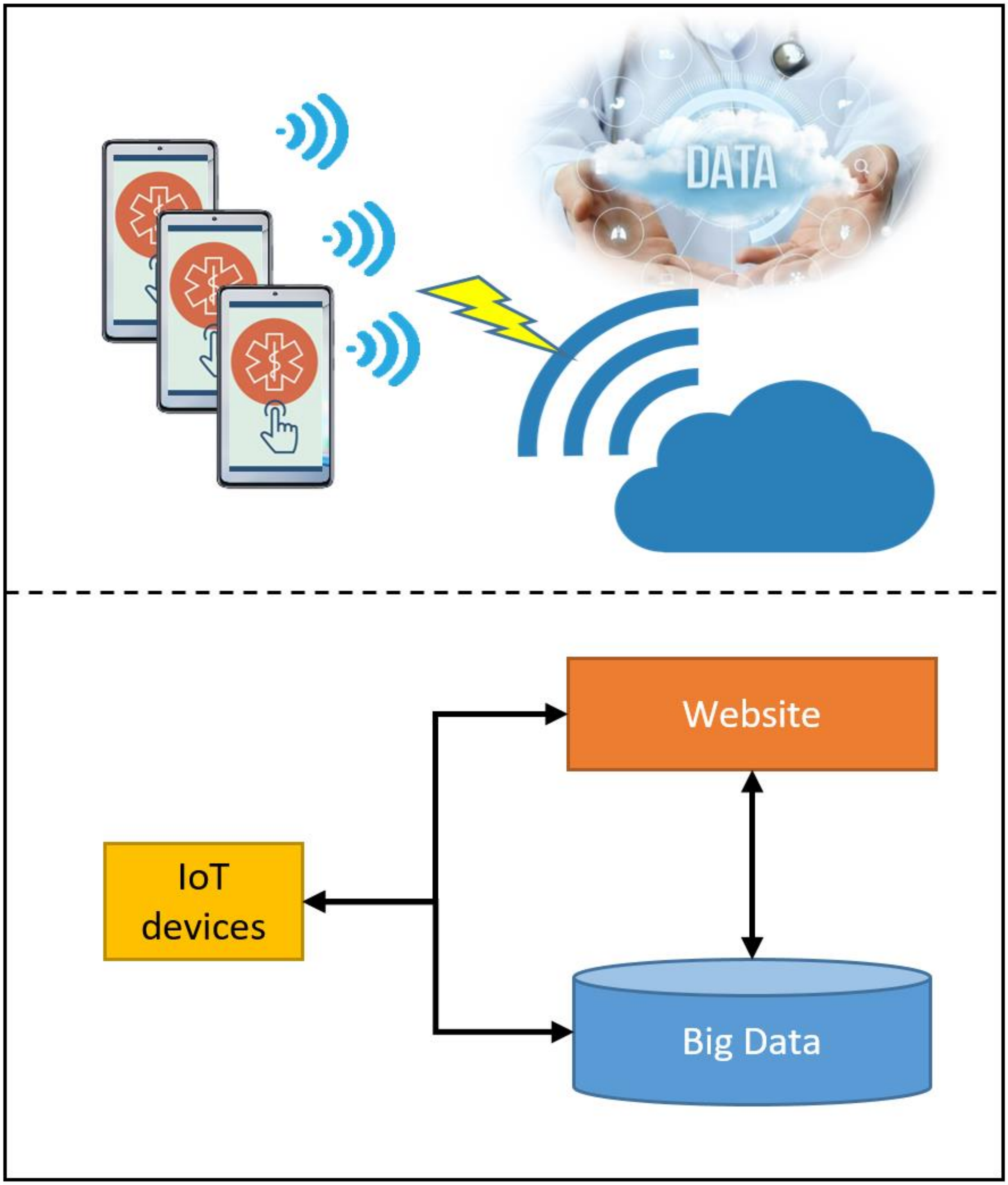

Figure 2.1 Concept of mobile edge for big data interaction.

Smartphones are also the main way people connect to the world and interact with it. A way to harness the power of the smartphones and help out patients is to connect patients to a cloud system via a mobile application. Connecting patients to a cloud system creates a "mobile edge" around the system. The mobile edge plays two important roles. First, it is the source of important raw data that flows into the cloud system. Second, it is what connects patients to their health information and their doctors. That motivates us to create a mobile application to connect patients with their doctors and data in a simple yet effective way to improve people's healthcare outcomes.

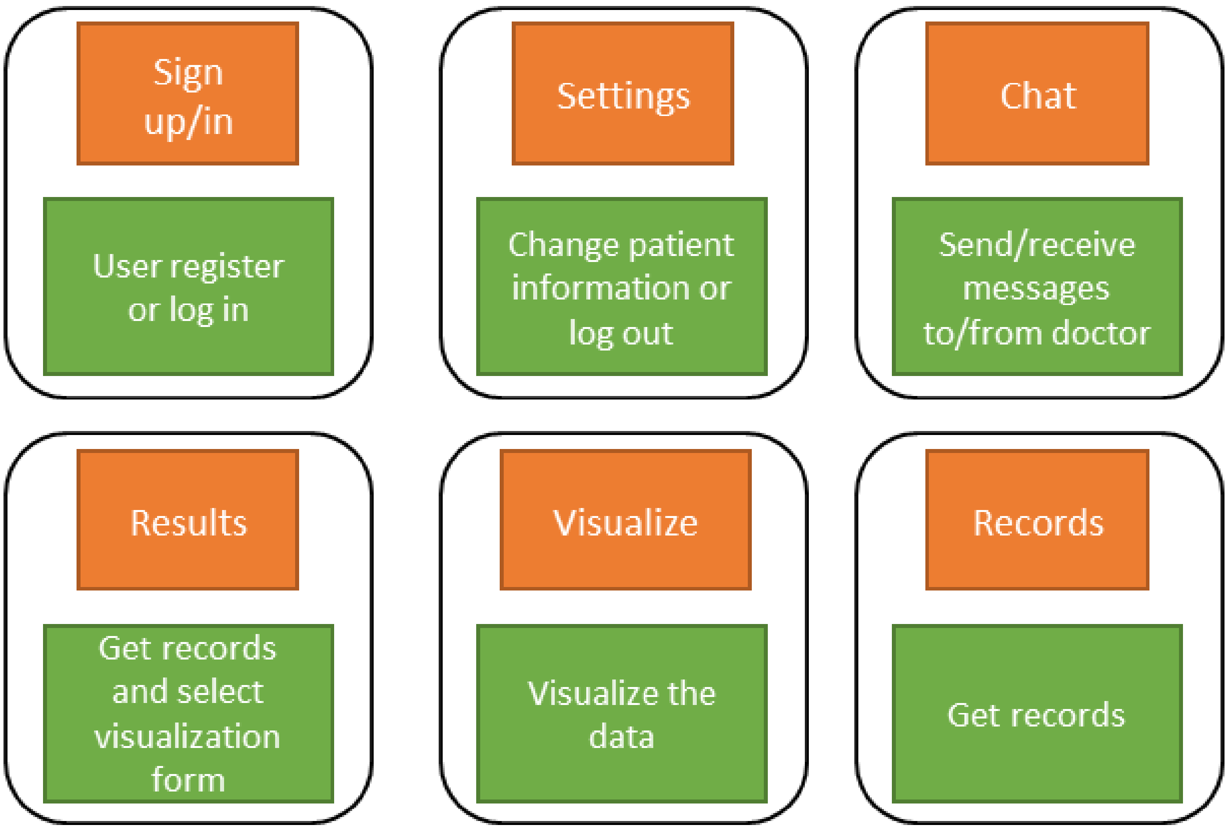

Figure 2.2 System Overview of the proposed s^2Edge system for smart and secure big data interaction on the mobile edge.

There are some previously reported mobile systems with healthcare purposes, but some limitations exist [17, 18]. Some studies reported mobile applications (apps) that can visualize the real-time data for the users [14, 19]. However, to further boost the potential of big data, it is important to stream the data to the cloud in a smart and secure way. So far, it is still pressing to develop a system that can effectively secure the data transmission. Further, it is important to provide an interactive interface that facilitates the communication between the edge user (like a patient) and the cloud user (like a doctor). Besides, the research in edge user information management is very limited [20, 21]. How to manage edge user information in the cloud and how to retrieve the information, are both important considerations of an effective edge system.

Focusing on the above challenges, in this study, we propose an innovative mobile edge system for smart and secure big data interaction, called s^2Edge where s^2 denotes Smart and Secure. This mobile edge system can be used not only for smart health big data, but also potentially for many other smart home and world big data applications. Here we take special interest in the precision medicine big data application.

s^2Edge can obtain essential physical activity data from the patient and send it to the big data cloud. The patient can use the s^2Edge to view the data that is being monitored and gain insights into their physical activities, which are related to many health concerns like aging, rehabilitation, and mobility difficulties. Also, the patient can directly converse with their doctor so they can be quickly informed and updated with any new information the doctor has learned from their data. All of a patient's data is secured behind a set of credentials that are managed by AWS. The mobile application is connected to an AWS cloud system. It is equipped to communicate with this cloud system in order to directly connect users to their health data and their doctors. It can also visualize real-time patient data to the system.

Overall, the proposed s^2Edge system is expected to enable users to access their data and ensure that these applications are smart and secure.

Our major contributions include:

(1) Big data visualization on s^2Edge for patient-phone interaction;

(2) Direct connection between doctor and patient via chat for patient-doctor interaction;

(3) Comprehensive patient management and big data communication for secure interaction.

The resulting mobile edge system is promising and is expected to demonstrate the potential of smart and secure big data interaction.

2.2 Mobile Edge for Smart and Secure Big Data Interaction Methods

2.2.1 System Overview

The purpose of the s^2Edge system is to provide a smart and secure interface for patients to get full access to their health data, receive real-time illustration of their physical activities, and communicate with their doctors. The system diagram of this mobile edge system is given in Figure 2.2. Each screen has a specific functionality that helps carry out the purpose stated above. First and foremost, data security is an important aspect of the system. In order to keep patient data safe, AWS Cognito is used to store and verify user credentials. Every patient must have login credentials to access their data, so that no unauthorized person can access their data. Data is not shared and there is no way to access another patient's data without accessing their account or being their doctor. Each patient has a doctor that they work with. In order to maintain and improve the health of their patients, doctors will create health records to track a specific aspect of a patient's health. These records and their associated analysis are made readily available via the Records and Results screens. These screens list out the health records associated with the patient. The Results screen will allow patients to view the associated activity data. Real-time analysis is conducted by retrieving phone sensor data and sending it to the cloud system. Upon starting up the real-time analysis feature, a direct connection is made to the AWS cloud system and data is captured from the phone’s sensors and sent to the cloud system. This feature is only activated when the patient turns on a sensor in the Visualize Screen. The patient always has the option to turn off phone sensor data capture. Not all sensors available on the phone will be captured. Only the necessary sensors that are listed by the doctor are used. The Visualize Screen will show time-series graph of each sensor output that is being used. Additionally, a Chat feature allows the patient to directly send messages to their doctor. This gives patients direct access to their doctors. The Settings screen is for patients to view and update their information. Patients can change the information they would like (e.g. name, height, weight, account login, etc.) as well as sign out of the mobile application.

2.2.2 Big Data Security

Security has become increasingly important, as many companies and users have found their data compromised. Since personal data and information is very important, it should not be able to be accessed by unauthorized users. In order to combat security breaches, s^2Edge utilizes AWS

Cognito [22] to safely store and verify user credentials. Cognito supports authentication, authorization, and user management for edge applications. Every patient has a unique identifier that allows them to access their data. This identifier is used throughout the system in order to retrieve data for a specific patient only. To protect patient data each user has a unique login so that no unauthorized user can access their account. This system design only allows patients and their doctor to have access to the data being collected. Additionally, each patient that signs up with the mobile application has their credentials securely saved in AWS Cognito. The login and signup screens are shown in Figure 2.3.

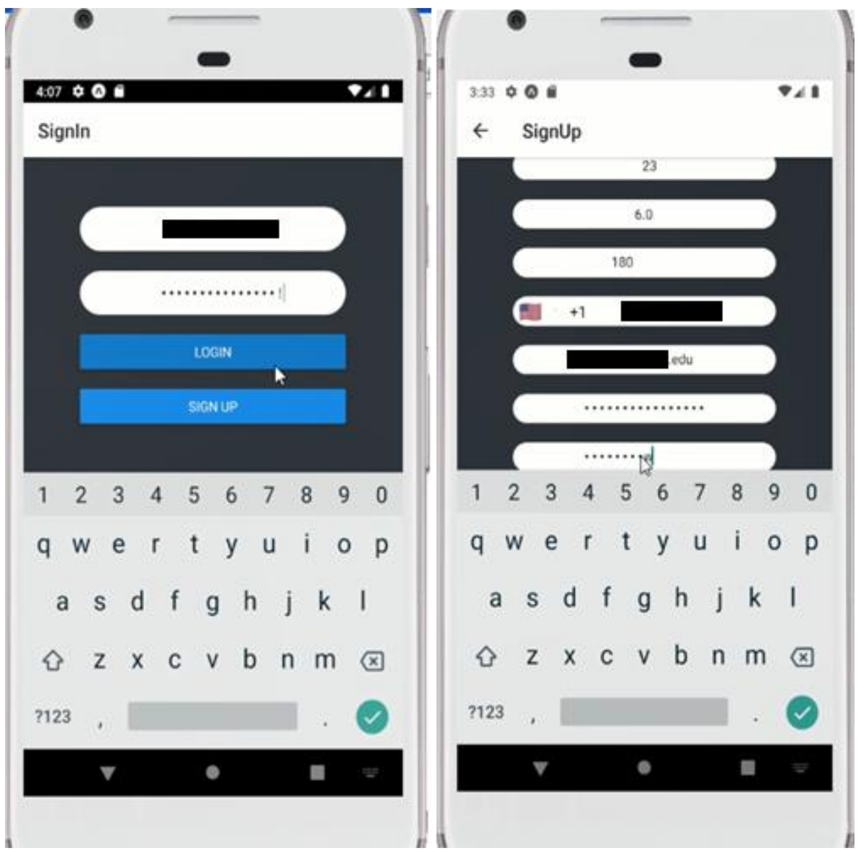

Figure 2.3 The user sign-in and sign-up pages, to secure the big data transmission and management.

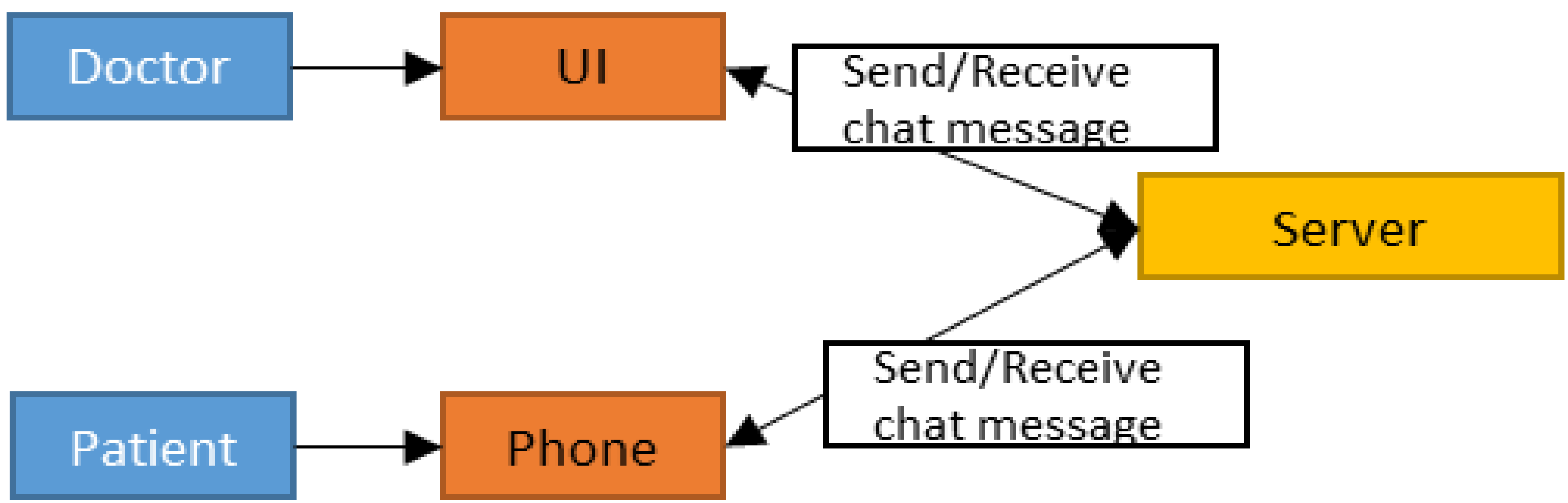

Figure 2.4 Patient-doctor interaction supported by the developed chat feature.

2.2.3 Interactive Chat

Doctor-patient communication is necessary for successfully improving and monitoring the patient's healthcare. We have developed a Chat feature that allows the patient to directly send and receive messages to and from their doctor. This gives patients direct access to their doctors. All messages are stored in the cloud so that no messages are ever lost so that all conversation history can be viewed. This means that doctors can quickly update their patients on any new feedback with their health, and patients can quickly receive replies from their doctor when they need medical suggestions. Figure 2.4 shows a basic diagram of how the doctor and patient communicate. Both the doctor and patient send messages through the server.

2.2.4 Record

Each patient has a doctor that they work with. In order to maintain and improve the health of their patients, doctors will create health records to track a specific aspect of a patient's health. These records are made readily available via the Records Screen, as shown in Figure 2.5. Each record contains data about the patient’s disease, a description about the disease, the start and end times of this record, the analysis being performed, and the sensors whose data will need to be captured on the phone. This helps patients understand what exactly is being analyzed so that there is full transparency between the doctor and patient. In addition, the patient has access to the past data. Going to the Results screen, the patient can select which record data they want to view. They can then further select which sensor data they want to view and on which date. Once the patient submits the request, the server processes the request and retrieves the archived data. Once the data is returned to the mobile app, a time-series graph is displayed showing the data that was captured.

2.2.5 Big Data Visualization

There are two main types of visualization that are available on the phone. The first is real-time visualization. Real-time analysis is conducted by retrieving phone sensor data and sending it to the cloud system. This feature is only activated when the patient turns on a sensor in the Visualize Screen. Patients have the ability to turn off the sensor data capture at any time. Only the sensors that their doctor selects will be used to capture data. The second type of visualization is historical. As real-time data is collected and analyzed, patients may desire to view their past data.

This data can be viewed via the Results screens. This screen lists out health records associated with the patient and gives them the ability to view the activity data associated with a particular record. The patient will be given the option of what sensor data to view and from which date. Once the patient submits a request to view the data, the data is retrieved from the cloud and visualized in a graph that the patient can cycle through and examine closely.

2.3 Mobile Edge for Smart and Secure Big Data Interaction Results

2.3.1 Experimental Setup

Experiments have been conducted to test the functions of s^2Edge. Before fully testing the functions of the system, test patients were signed up with AWS Cognito. After signing up the test accounts, the following were tested: the test accounts could login to the mobile app; test accounts could view and update their information; test account information was saved on the AWS cloud system; test accounts could visualize sensor data; test accounts could visualize historical sensor data; test accounts could chat with their doctor.

2.3.2 Record

Upon startup of the mobile application, the patient is prompted to login or signup. The process of authorizing patients during logging in or signing up is managed by AWS Cognito. The mobile application also gives the ability to update patient information and their credentials. Once a patient is fully signed up with the system, a doctor can connect with the patient and start helping them analyze their health.

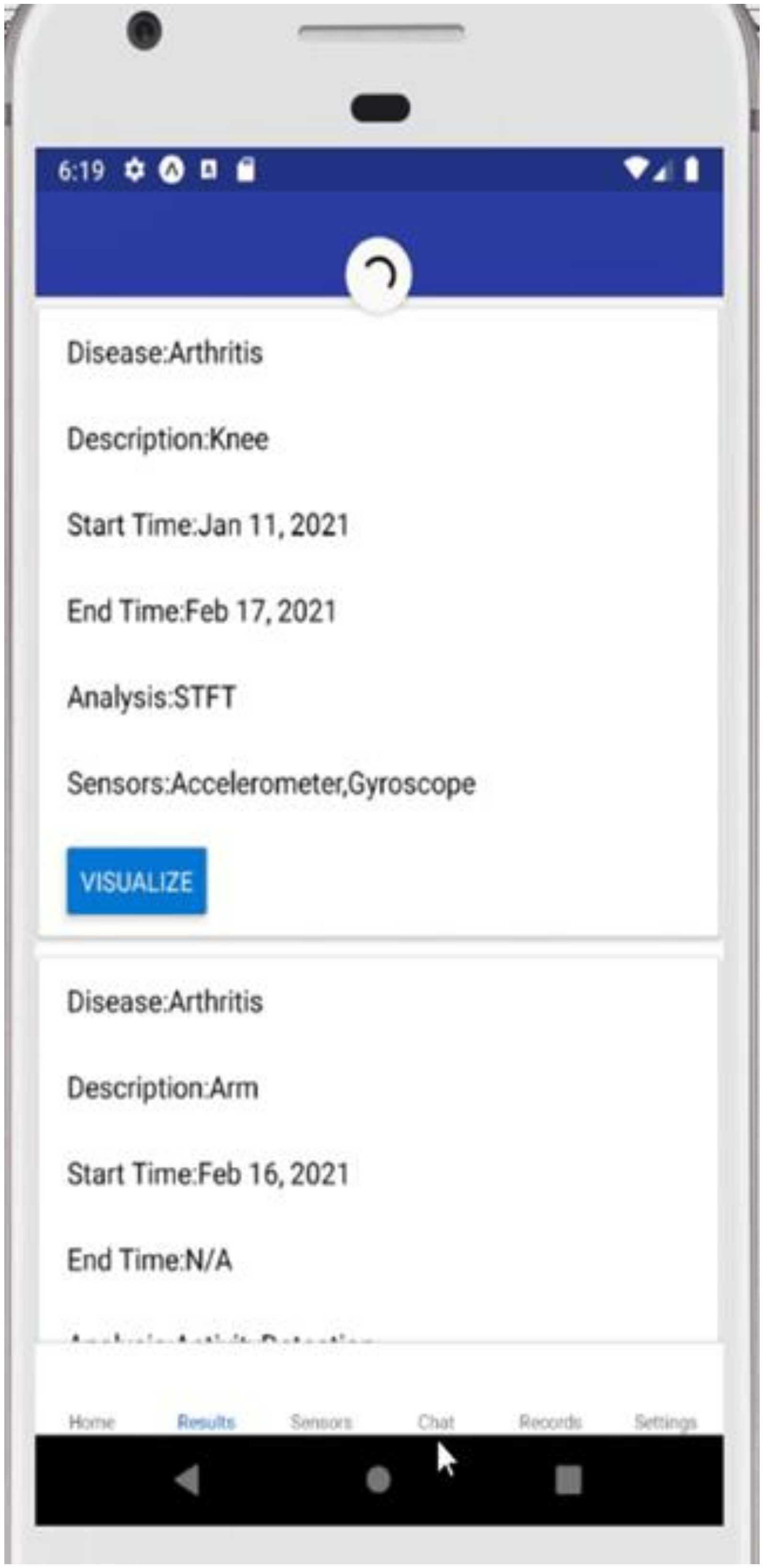

Figure 2.5 Patient record for big data management.

Figure 2.6 Patient-doctor interaction through instantaneous chat.

The doctor can create records associated with the patient that will help track the patient's health. These health records contain information about what aspect of the patient's health is being tracked and what analysis is to be performed with the patient's smartphone data. The records also have the types of sensors that will be used for analysis. The patient's health records and their associated data can be found in the Results and Records screens on the mobile app. The Results section is the main screen that will allow patients to connect to the activity data that their smartphone is producing. The patient selects the particular record to view and the specific sensor activity data. This functionality was tested by having a doctor account create records for the patient. The records were then viewed on the Results and Records screens.

2.3.3 Chat

To test the Chat feature of s^2Edge, messages were sent between the doctor and patient. Correct messages were displayed on the Chat screen as shown in Figure 2.6, and the messages were saved on the AWS cloud system. When testing this aspect of the system, the doctor account is also setup so that messages can be sent back and forth between the doctor and the patient. In order to communicate with their doctor, patients simply type a message into the message box and press the submit button. This message is instantly saved and is available for the doctor's viewing. In the same way, doctors can view their conversation with their patients and send messages back. This is a significant result because it demonstrates the direct connection between a doctor and their patient. This means that patients can have direct access to their doctors and constantly communicate. No longer will patients feel disconnected from their doctors and feel unsure about their health. Now doctors can quickly inform patients of new information on their health and help them improve.

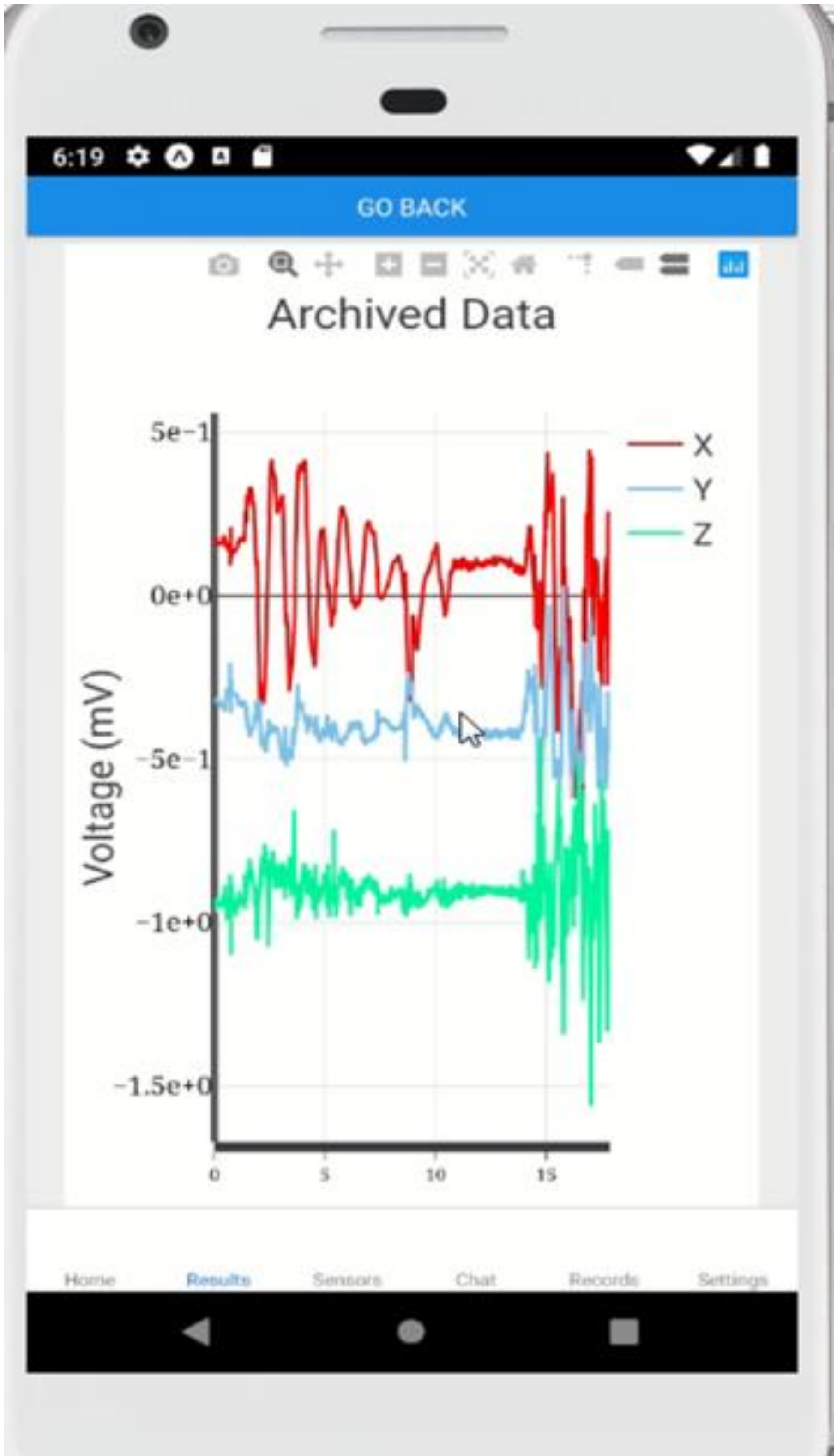

Figure 2.7 Patient-edge interaction through visualization of patient data.

2.3.4 Visualization on the Edge

To test viewing the historical data sent by the patient, the Results screen was selected and a request was submitted to view data. It was then shown that the corresponding graph was created when the data was returned from the AWS cloud system as shown in Figure 2.7. When testing this aspect of the system, there already existed past data that has been sent when the patient was sending real-time data. This part of the mobile application allows patients to look at their past history of activity data.

When the patient wants to view this data, they must select the record they want to view and the specific sensor data. When this request is created and sent to the cloud system. The cloud system then processes this request and returns the data to the edge. A time-series graph is created to visualize the stored data. Note that not all of the data is viewable at one time. This is because it allows the patient to get a closer look of each segment of data and there is most likely a lot of data that is stored on the cloud for a given date. Since there is potentially a lot of data, it might not be possible to retrieve and visualize the data all at once. Some barriers include request timeouts and responses that are too large.

To address these problems, only a set interval of the data is retrieved at a time. The retrieved data is appended to the existing data in the graph and only a new request is sent when the patient attempts to view past the current data in the graph. This ensures that data does not take too long to retrieve and it ensures that not too much data is sent back to the patient. The patient can cycle through the time-series graph by pressing the “Forward” or “Backward” buttons listed above the graph.

When the patient is done looking at their data, they can return back to the main results screen by pressing the back button listed at the top. This is a significant feature and result because it demonstrates that the patient is able to directly interact with and view their big data. This means that patients are involved in the health monitoring process as well.

Figure 2.8 Big Data Co-visualization on Edge-Cloud, indicating effectiveness of proposed s^{2}Edge system and how it works together with cloud.

2.3.5 Big Data Co-visualization on Edge-Cloud

In order to test the real-time functionality, it was verified that on the Visualize screen the correct sensor graphs were displayed and that the data was flowing through the AWS cloud system. The data was also viewable on the doctor's account in real-time. When testing this aspect of the system, both the mobile application and website were pulled up. As shown on the left side of the Figure 2.8, the mobile application visualizes the sensor data on a time-series graph. Different channel axis values are being captured from the sensors and graphed. All of the data is sent in real-time to the cloud system. The mobile application makes a direct connection to the cloud system and is able to capture and send data directly to the cloud.

The doctor, who uses the website on the right, can subscribe to the data that the patient is sending and view it in real-time. Notice in the figure, the doctor's graph and the patient's graph are very close to in-sync. This means that the doctor gets almost instantaneous updates of what the patient is currently doing and is able to get an in-depth analysis of the data. The graphs only show a certain interval of time, so as more data comes in, the graphs will shift so that new data can come in. Notice also that each graph is equipped with controls that allow the patient and doctors to look at only certain axes or increase/decrease the view size. This enables both the doctor and patient to get a close look at the data they want. This is a significant result because it makes it easy for patients to be monitored with a touch of a button. These results are also significant because it allows doctors to better monitor and analyze their patients. Doctors can now get real-time information on their patients so they are better able to analyze and improve their healthcare.

2.4 Conclusion

Mobile edge applications play a vital role in big data interaction, and in this study, we propose an innovative system, called s^{2}Edge, to enable smart and secure big data interaction on the mobile edge. This edge system, by leveraging AWS security features, can provide a secure way for patient registration and management, as well as data protection. Furthermore, the system has a highly interactive capability that enables users (like patients) to access their data and records in addition to communicating with the cloud users (like physicians). The resulting mobile edge system is promising and is expected to demonstrate the potential of smart and secure big data interaction. s^{2}Edge is expected to, not only demonstrate the potential of smart and secure big data interaction, but also show the potential for many emerging smart home and world big data applications.

3. A NOVEL CLOUD SYSTEM FOR SMART AND SECURE BIG DATA

The promising era of big data is advancing many new practices. One of the emerging big data applications is healthcare, which requires smart and secure transmission and management of the big data, including both patient information and patient measurements. This study aims to demonstrate a novel big data cloud system, s^2Cloud, which can help enhance healthcare systems to better monitor patients and give doctors critical insights into their patients' health. s^2Cloud achieves big data security through secure sign up and log in for the doctors, as well as data transmission protection. The system allows the doctors to management both patients and their records effectively. The doctors can add and edit the patient and record information through the interactive website. Furthermore, the system supports both real-time and historical modes for big data management. Therefore, the patient measurement information can, not only be visualized and demonstrated in real-time, but also be retrieved for further analysis. The smart website allows doctors and patients to interact with each other effectively through instantaneous chat. Overall, the proposed s^2Cloud system, empowered by smart secure design innovations, has demonstrated the feasibility and potential for healthcare big data applications. This study will further broadly benefit and advance other smart home and world big data applications.

3.1 Introduction

The world is now defined by mass production and consumption of data. The amount of data generated by the world continues to increase exponentially each year and shows no signs of stopping. It is predicted that the annual size of the global datasphere will increase to 175 ZB in 2025 [23]. We rely more heavily on this data to make decisions and to make our world smart [24-27]. Due to these increases, new challenges arise to manage the massive amount of data, which is known as “big data”. As a result, systems from different industries must now adapt to deal with these challenges. While new challenges appear, however, technology to manage this data driven world have been developed and are continually improving. Big data consists of multiple characteristics like volume, velocity, and variety. These characteristics imply that multi-type large amounts of data are quickly generated. It is pressing to develop innovative and effective big data systems to manage big data.

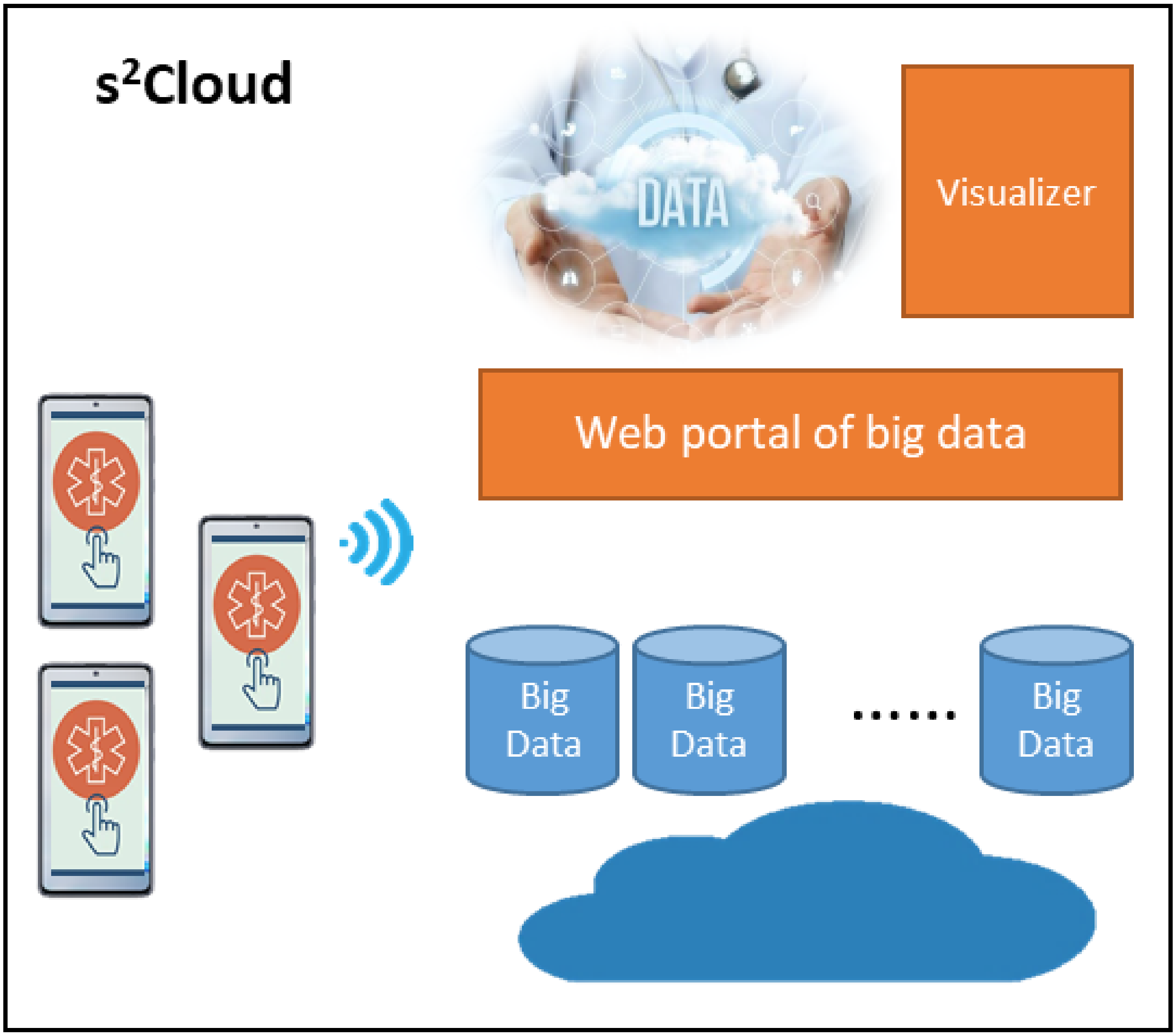

Figure 3.1 Smart and secure big data management system for s^2Cloud application.

We take special interest in the healthcare big data area [28-30], and research how to tackle these challenges through innovative design as shown in Figure 3.1. The technologies can also be generalized to many other smart home and world areas. Good health is a very important aspect of each person's life and therefore it is always worth to strive to help improve and invent ways to better people's health. In this study, we investigate how to enable smart and secure big data management on the cloud as shown in Figure 3.1, thereby advancing big data healthcare.

There are some previously reported studies for healthcare data capturing and management. Some studies have mainly focused on the data capturing using wearable sensors [31-33] and have not developed the cloud system for big data management. However, demonstrating and storing big data in the cloud is highly important for medical decision support [34-36]. Besides, some other studies have not fully implemented the big data management features [37-40], such as secure doctor and patient management, interactive chatting, and visualization functions. They lack a system that can effectively, securely, and in a smart way, manage the big data.

To enable effective, smart, and secure big data management, we propose an innovative system, called s^2Cloud. s^2Cloud achieves big data security through secure sign up and log in for the doctors, as well as data transmission protection. The system allows the doctors to manage both patients and their records effectively. The doctors can add and edit the patient and record information through the interactive website. Furthermore, the system supports both real-time and historical modes for big data management. Therefore, the patient measurement information can, not only be visualized and demonstrated in real-time, but also be retrieved for further analysis. The smart website also allows doctors and patients to interact with each other effectively through instantaneous chat. s^2Cloud, implemented on Amazon Web Service (AWS) [41], can help enhance healthcare systems to better monitor patients and give doctors critical insights into their patients' health.

Our major contributions include:

(1) Design a secure big data management system, with doctor sign up and log in functions, as well as data transmission protection;

(2) Design an interactive doctor-cloud interface, which allows the doctors to easily access the data, and communicate with the patients via chatting;

(3) The system allows the doctors to conveniently manage both patient information and patient records;

(4) The system supports both real-time and historical demonstration of the big data.

Overall, the proposed s^2Cloud system, empowered by smart and secure design innovations, has demonstrated the feasibility and potential for healthcare big data applications. This study will further broadly benefit and advance other smart home and world big data applications.

3.2 Methods

3.2.1 System Overview

The proposed system can provide a smart and secure big data system for doctors to access to their patients' health data, better manage and understand the data, and communicate with the patients. The system diagram is shown in Figure 3.2. In order to keep doctor data safe, verification of user credentials was done by AWS Cognito. Login credentials are required for every doctor to access their data. This is done to ensure no unauthorized user has access to critical data. Data is not shared, and a doctor's data is only accessible using their account credentials. Doctors can use the Patient Management page to add patients to their patient list. Doctors can create health records to help organize, track, and improve their patients' health. Health records and their associated analysis are viewable via the Records page. This page lists out all the health records associated with each patient. The Dashboard page will allow doctors to view the patient data. Real-time visualization can be used to view the real-time data coming from the patient's phone. In addition, historical visualization of past patient data can be viewed on the Dashboard page as well. The doctor can view past data from a specific patient. Another page that is available is the Chat page. The Chat page allows the doctor to directly send messages to their patients. This gives doctors direct access to their patients conveniently and vice versa. The website also has a Settings page. This page is for doctors to view and update their information. Doctors can change their information if they would like as well as sign out of the website.

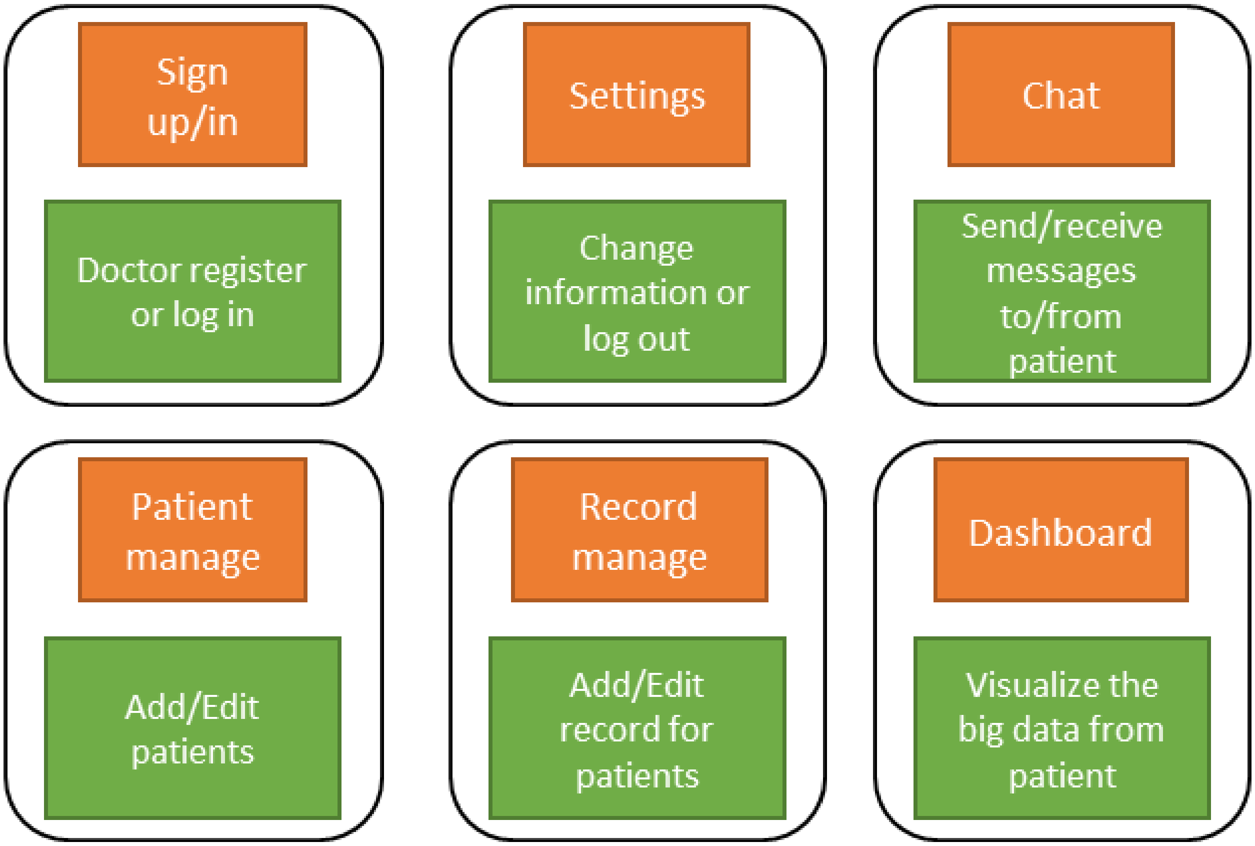

Figure 3.2 System diagram of the proposed s^{2}Cloud system, for smart and secure big data management.

3.2.2 Big Data Security

Security is a critical component of healthcare data is applications [42-46]. Here we have proposed a mobile-cloud collaboration mechanism to provide secure data access. We have introduced AWS Cognito to verify user credentials, through robust user authentication and authorization, for both cloud website and mobile APP users. Doctors have a unique identification number and login credentials that are used to help retrieve their data. These security measures help ensure no unauthorized person can access important data. Data is not shared with anyone and access to it is only possible by using their account credentials.

Additionally, each doctor that signs up with the website has their credentials securely saved in AWS Cognito. The processes of signing up and signing in is shown in Figure 3.3. Signing up involves filling out doctor information (e.g. name, email, hospital, etc.). Once this data is filled out, a request is sent to AWS Cognito where the account is created with the credentials the doctor entered. Then a verification code is sent to the doctor's email. The doctor must enter in the verification code to the website so that their email is verified.

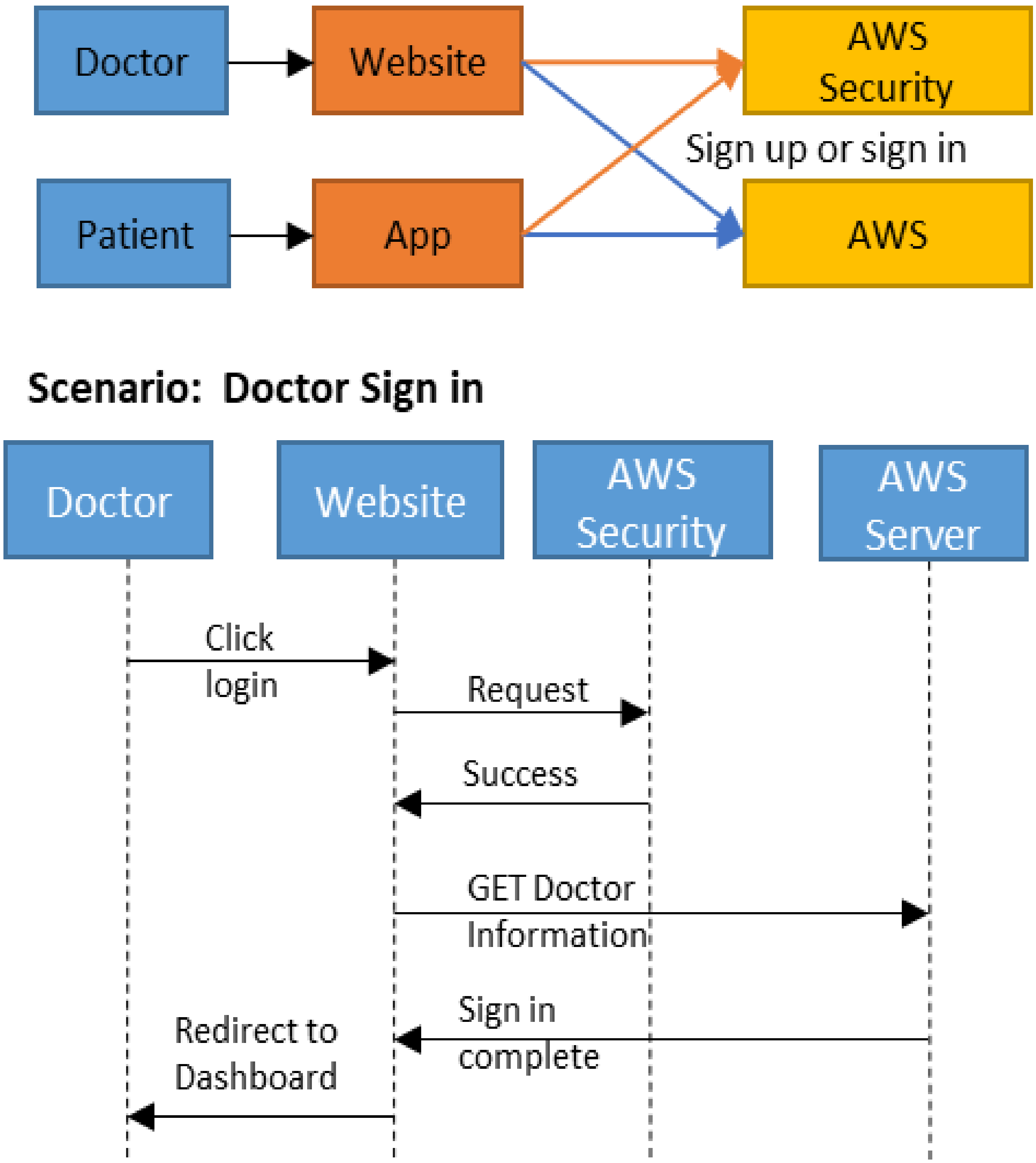

Figure 3.3 Doctor sign up or sign in flow and protocols.

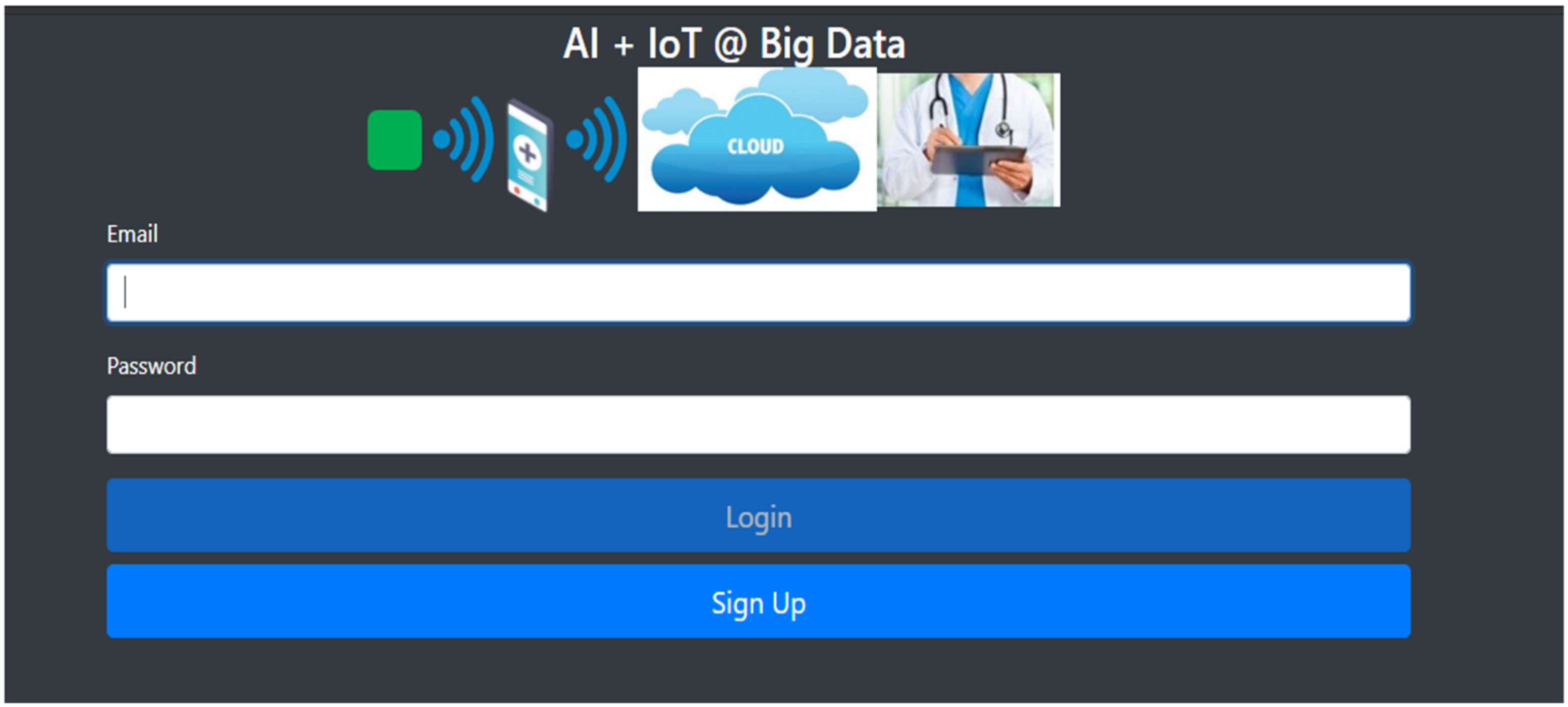

Figure 3.4 Sign up or log in page of the s^2Cloud system.

The login and signup screens are also shown in Figure 3.4. Upon signing in via the website, a sign-in request is made to AWS Cognito. If this request returns that the doctor credentials are valid, an additional request to retrieve that valid doctor's information is made. Retrieving this information furthers security of the system because the information is used in requests so that only valid doctors can make valid requests.

3.2.3 Interactive Chat

In big data applications, especially healthcare, effective doctor-patient communication is critical. That is why the website is enabled with a chat feature that allows the doctor to directly communicate with their patients. Doctors can use the chat feature to quickly notify their patients of new information regarding their health.

Likewise, patients can instantly receive replies from their doctor about any questions or concerns they have regarding their health. Doctors and patients use the chat feature by typing out a message and pressing the submit button. All messages that are sent go through the server and saved on the cloud. Messages are available for viewing whenever the conversation is opened up. When the conversation is rendered, the data associated with each message (e.g. sender, date, content) are displayed. If the sender was the doctor, then the message is colored blue. If it was the patient, then it is colored red.

3.2.4 Patient Management

Every doctor has multiple patients that are under their care. The Patient Management page gives a list of all of the current patients that they are helping and their associated data (e.g. patient id, height, weight, etc.). In this page, doctors have the ability to add available patients under their care by pressing the 'add patient' button.

This button will open up an 'add patient' window that will allow the doctor to select a patient that does not yet have a doctor. On this page, the doctor can edit any of their patient's information like their height, weight, etc. as shown in Figure 3.5. Finally, doctors also have the ability to remove patients from under their care.

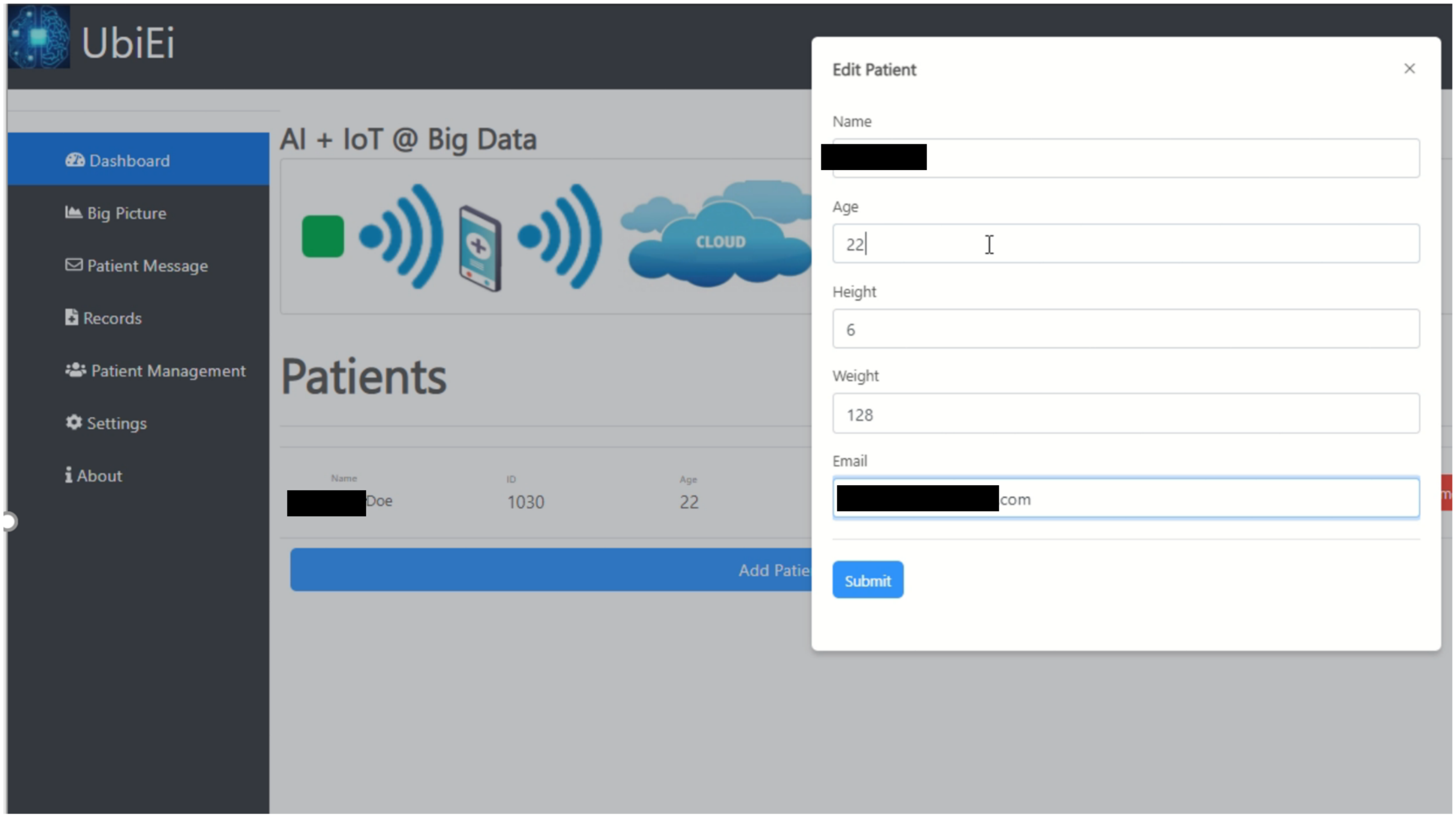

Figure 3.5 Add or edit patient information for a doctor.

3.2.5 Record Management

Doctors are associated with many patients. To keep track of a patient's health, doctors can create health records. These health records are used to help monitor patients so that doctors can maintain and improve patient health. Patient records can be viewed using the Records page. A record contains the following information: disease type, record description, start and end times of the record, the analysis type performed, and the sensors to be used to capture data for the record. Electronic health records help doctors stay organized and let them quickly retrieve the data they need to know about their patient.

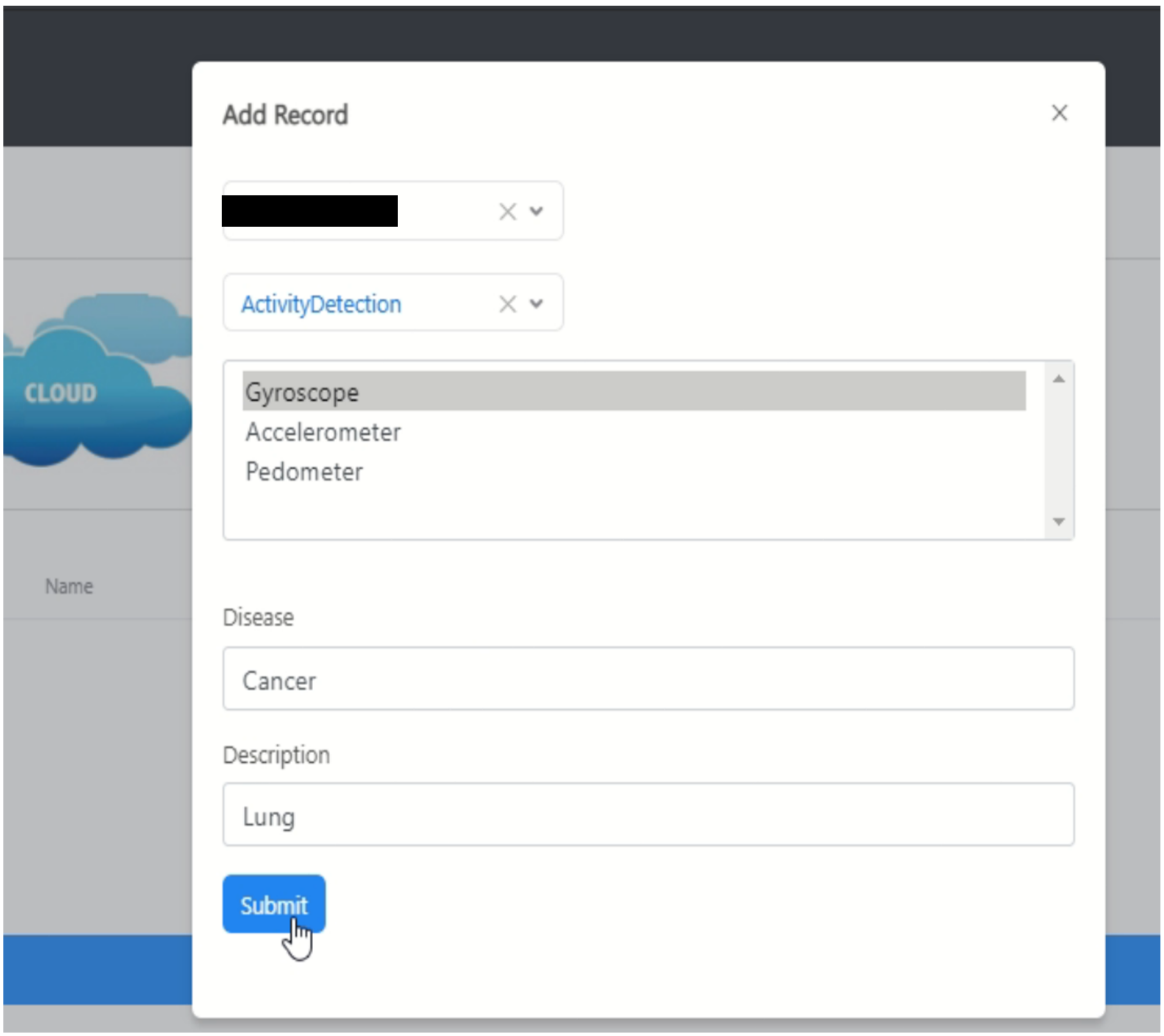

Figure 3.6 The 'add record' function for a patient.

Upon opening the page, a request is made to the server to retrieve all records of patients associated with the signed in doctor. Doctors can add as many new records they would like to a patient. Adding a record will bring up a separate window as shown in Figure 3.6. The doctor will need to fill out the different fields and then press the submit button. Once a doctor adds a record, that record will show they have the ability to edit it or remove it. If the doctor edits a record, a popup form similar to the 'add record' popup window will appear. The 'edit record' window will first display all of the current values for each property of the record. The doctor can then modify the properties of the record. Figure 3.7 shows the sequence of events that are involved with reading and editing records.

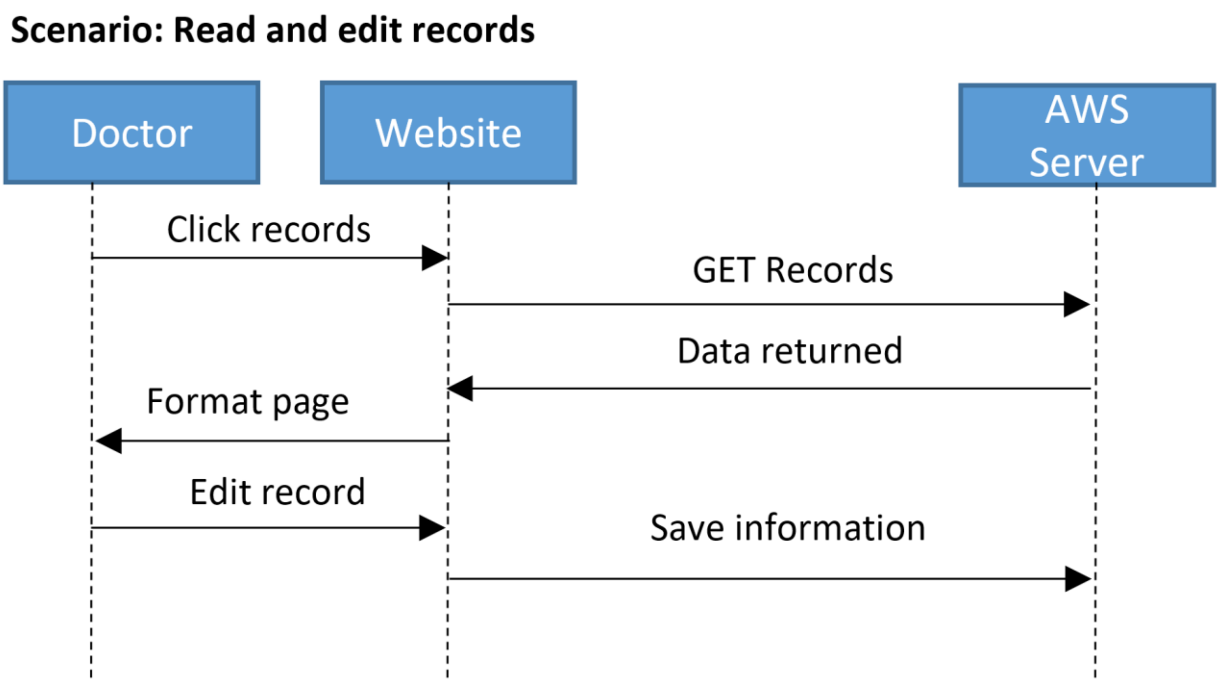

Figure 3.7 Read and edit record protocols.

3.2.6 Big Data Management

An important function of the website is accessing the big data stored on the cloud. This is done through the Dashboard page using the historical visualization button. Upon pressing the historical visualization button, a popup window is shown with several dropdowns. Before a doctor can retrieve stored data, they must first select the patient, the patient's record, one of the sensors that is on the patient's record, and the date the data was sent to the cloud. The upper part of Figure 3.8 shows that as the doctor makes these selections in the submission form, requests will be sent out to retrieve the corresponding data to populate the dropdowns for the following sections later on in the form. Once all of this information is selected, a request is submitted to retrieve the stored patient data. The request is taken in and the server returns the data. The data is finally returned to the doctor and the website creates graphs to visualize the retrieved data.

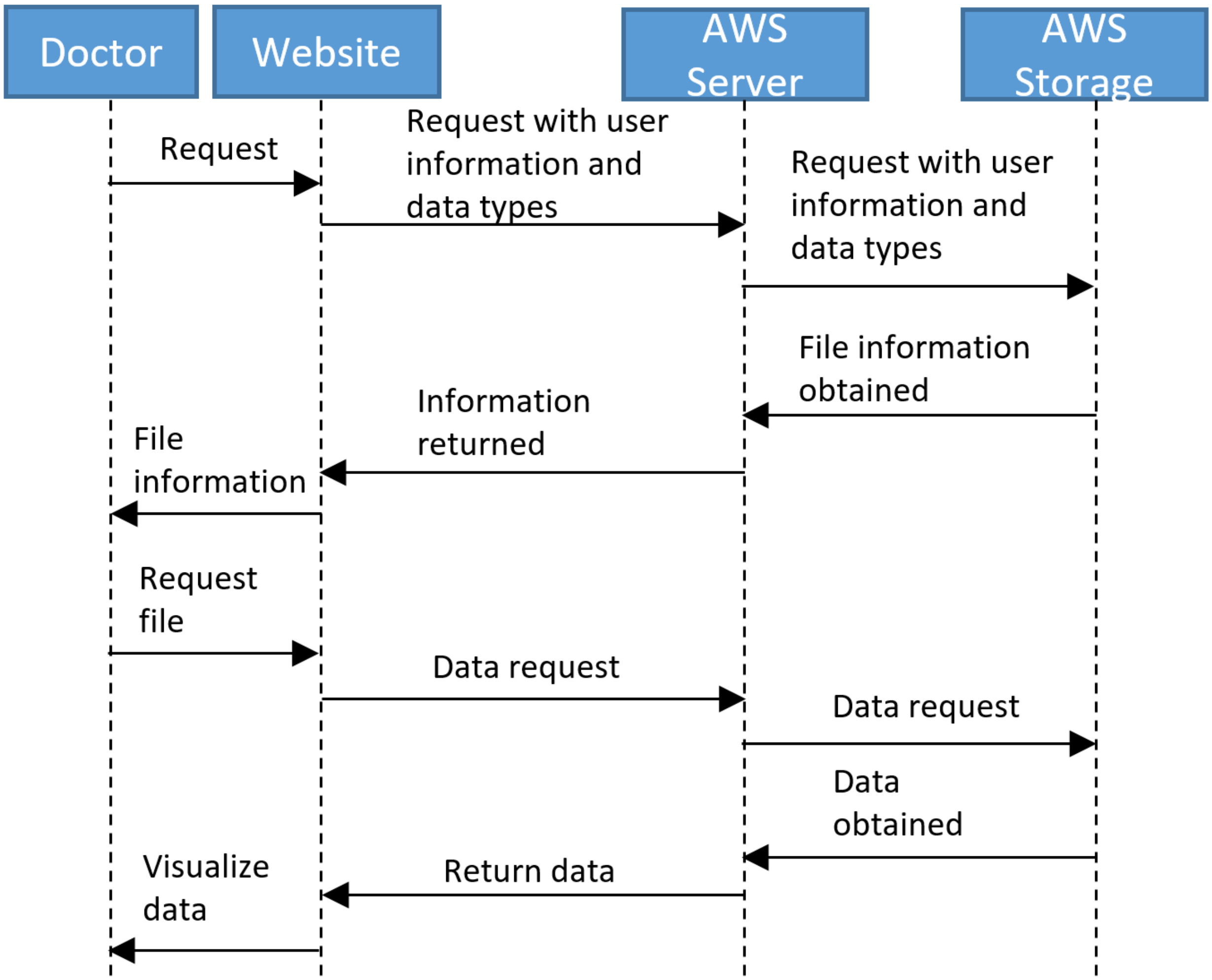

Figure 3.8 Big data access protocol.

3.2.7 Big Data Visualization

After a doctor retrieves the data from the cloud, the data is visualized using a time-series graph for the input data. The graphs created contain several functionalities that allow doctors to closely view the data such as zoom in/out. One complicated task associated with visualization is viewing large amounts of data at once. Since a large amount of data is in cloud storage, retrieval and visualization can be difficult. Some reasons for this include timeouts when retrieving the data and large responses. To address these problems, data is retrieved in set intervals only. This method is used so that requests to retrieve data do not timeout and that the response size is not too large. Since not all of the data is viewable at once, the doctor is given buttons (backward, forward, play, stop) that allow them to scroll through the data. If the doctor scrolls past the end of the dataset they have stored in their browser, then a request is made to retrieve more data. Each time a request is made, the data is appended to the data the doctor currently has already in their browser. These graph features allow the doctor to closely analyze all of a patient's data without being hindered by the size of the data.

3.3 Results

3.3.1 Experimental Setup

We have tested the developed functions on s^2Cloud to demonstrate the effectiveness of the proposed big data management system. In order to test the different parts of the system, test accounts have been created for the doctors and the patients using the respective signup pages. Afterwards, different pages with various functions have been tested.

3.3.2 Web Layout

The website consists of different pages that enable the doctor to fully manage and monitor their patients. The pages are setup such that the design of the page is simple and clear and the functionality is easy to use. Doctors navigate the website using the main side navigation bar on the left side of the page as shown in Figure 3.9. This navigation bar is available on all pages. Upon clicking on a different tab within the navigation bar, the corresponding page is loaded. Each page is identified with a unique URL. The components within each page consist mainly of buttons, lists, dropdowns, and forms.

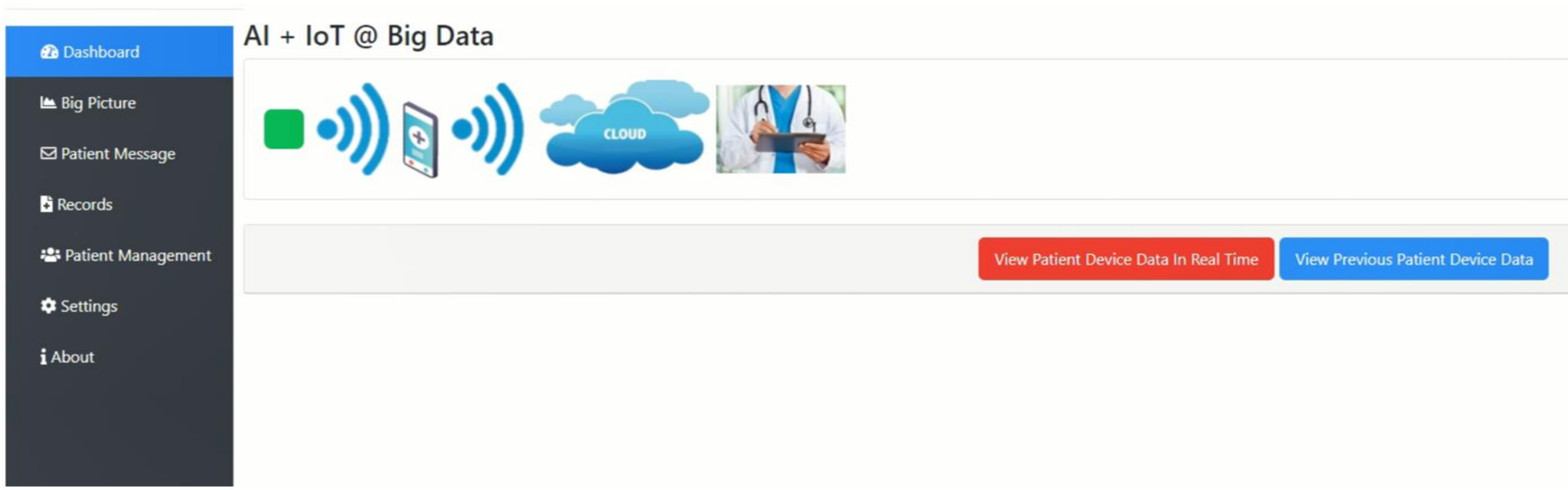

Figure 3.9 The developed big data website layout.

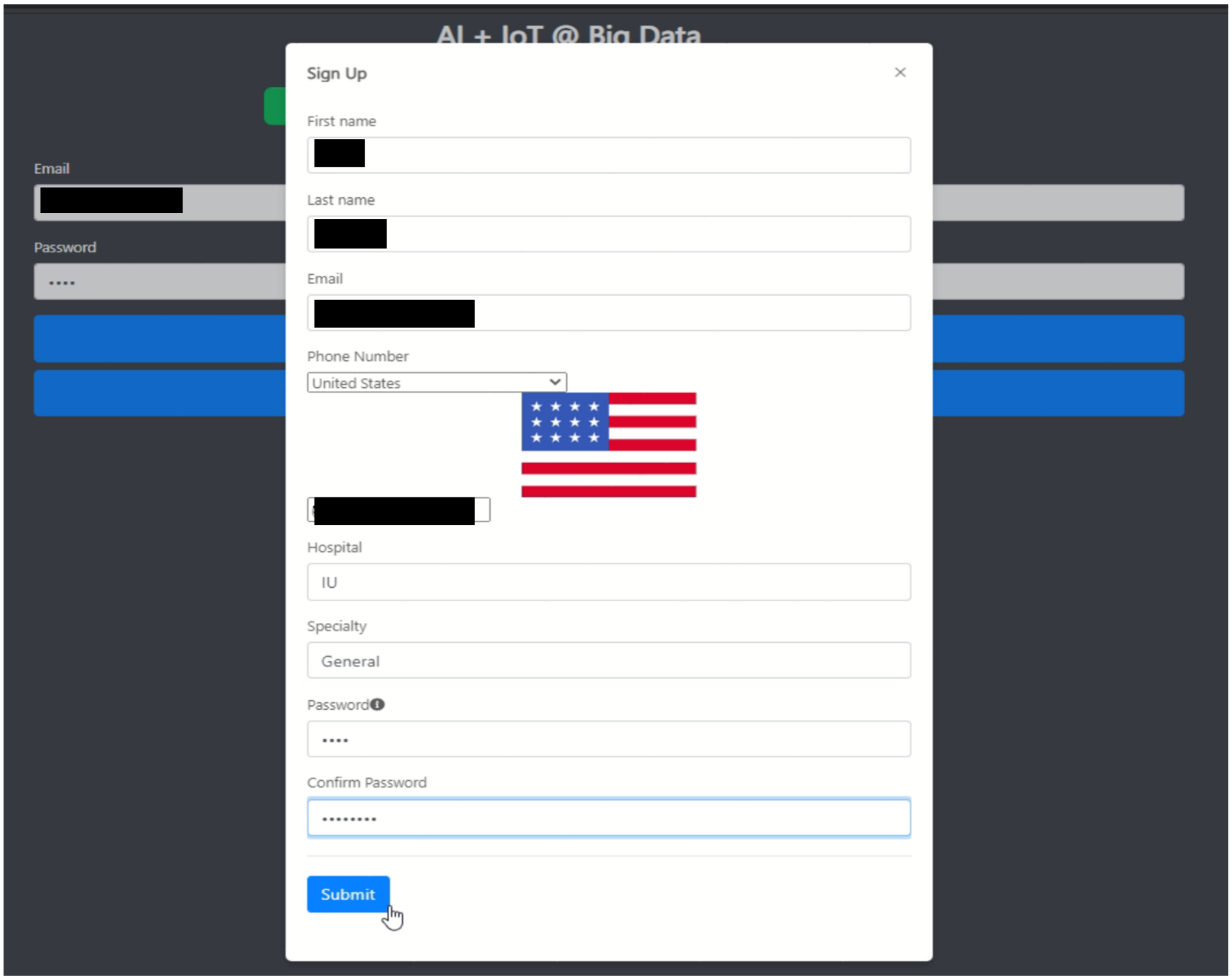

Figure 3.10 Doctor sign up page.

3.3.3 Sign Up/In

For the security testing, doctors were signed up with AWS Cognito. When a doctor attempts to sign up, they must fill out the necessary information shown in Figure 3.10. When the form is fully filled out, it is sent to AWS Cognito where then the doctor's credentials are saved. After this, an email containing a verification code is then sent to the doctor's email. The doctor must enter in the code to the verification code popup window to verify their email shown in Figure 3.11. After the doctor's email is verified, the doctor's information is then saved to the cloud and given a unique identifier.

Once this is complete, the doctor is redirected to the main dashboard page. When a doctor attempts to sign in, they must fill in the email and password fields. After the doctor presses the login button, the credentials are sent to AWS Cognito for verification. If the credentials match an existing account, a success is returned. If a success is returned, the doctor's information is then fetched from the cloud. Finally, the doctor is redirected to the main dashboard page. Only preexisting doctor accounts could sign in and that their information was fetched from the cloud.

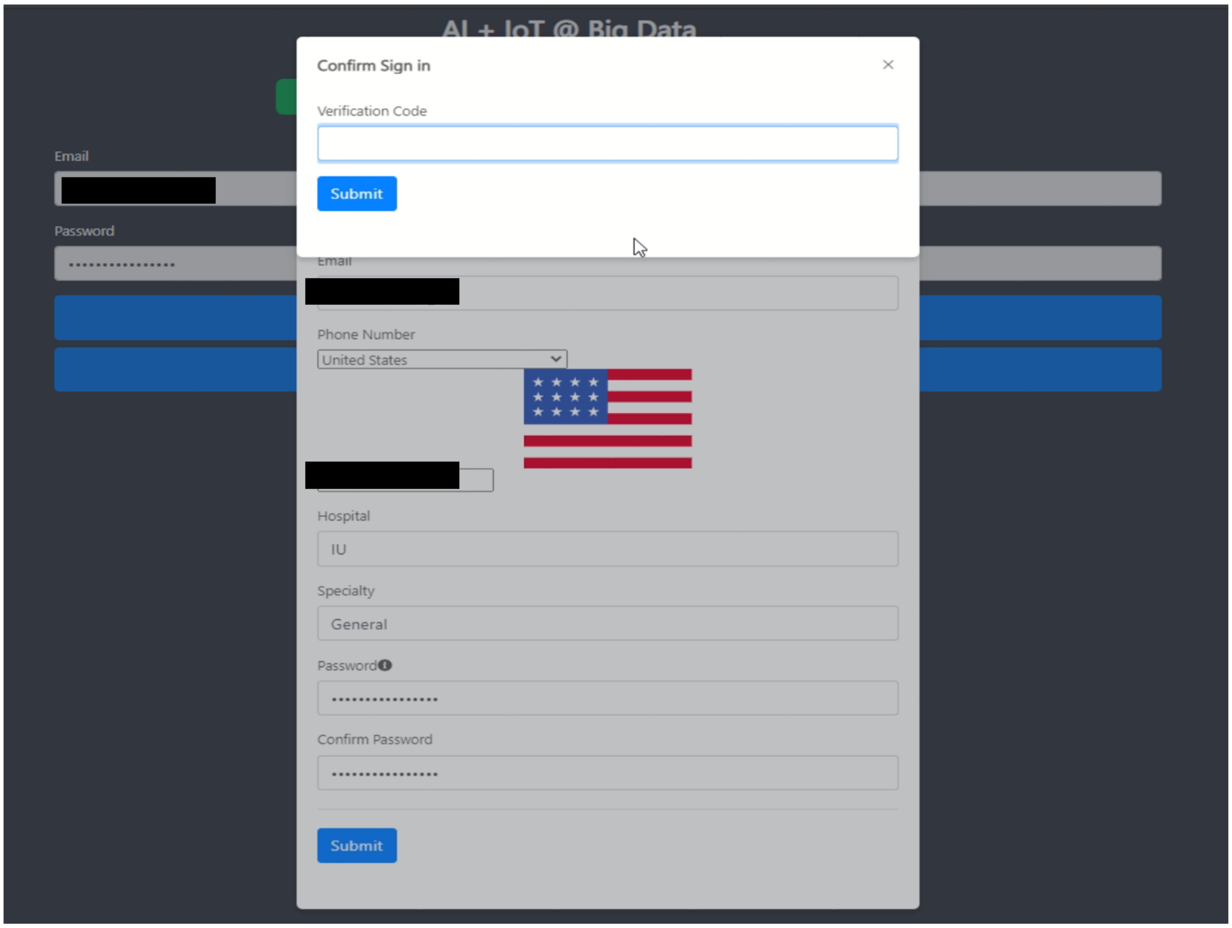

Figure 3.11 Verification page for sign up.

3.3.4 Patient Management

Figure 3.12 shows the added patients for a doctor. Detailed information like patient ID, age, height, weight, and email are demonstrated. The doctor can click the 'add patient' button to add more patients or click the edit or remove button to change the patient information or delete the patient information. With his patient management page, the doctor can easily go through and/or update the patient information if needed.

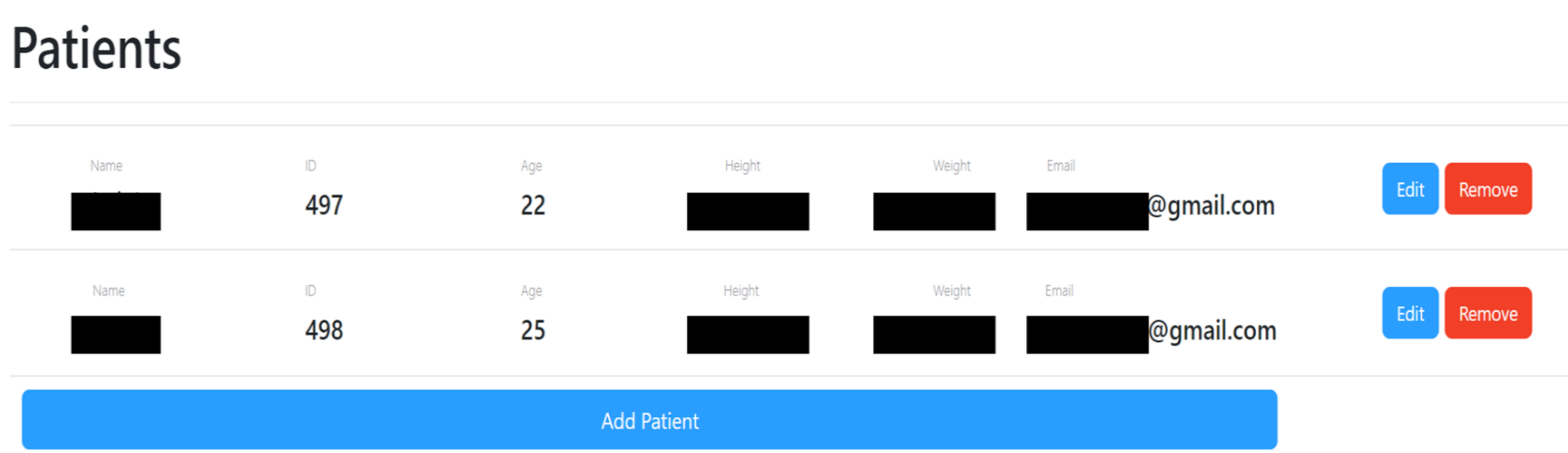

Figure 3.12 The patient information added.

3.3.5 Chat

Messages were exchanged between the test doctor and test patient to demonstrate the functionality, as shown in Figure 3.13. Since doctors have multiple patients, they must first select the patient they wish to converse with in the list on the left side of the page. The chat will be rendered on the right side along with a message box and submit button.

Doctors and patients communicate with each other by typing messages into a message box and sending them. Sent messages are saved to the cloud and are readily available for viewing. Each message has sender, date, and content information associated with it. When messages are rendered, this information is displayed. To help make conversations more readable, on the website, doctor messages are colored blue and patient messages are colored red. The chat feature enables doctors and patients to have a direct connection with each other. This is significant because patients are allowed to have direct access to their doctors and quickly communicate the information. Patients no longer have to feel uncertain about their health and detached from their doctors. Patients can be quickly informed about new information regarding their health from their doctor.

3.3.6 Big Data Storage and Visualization

Visualizing the big data stored on the cloud is critical for big data management. The doctor can either visualize real-time data or historical data. On the Dashboard page, there are two buttons that help the doctor to select the exact patient data they want to view. When the patient data is retrieved, a graph will show the time-series data from the sensor. For the real-time case, a test patient account was setup and was used to send test data. When data is sent from the patient, it is saved to the cloud storage in real-time.

A doctor account was used to visualize the real-time data sent from the patient. Before real-time data can be viewed, the patient must be actively sending data. The cloud system is used to keep track of whether patients are active or not. So, if a patient is not active then the doctor will not be able to view real-time data. If the patient is active then a real-time visualization request can be sent. The system is setup so that only authenticated users with Cognito can view the data. Once this occurs, the system determines the necessary information to view real-time data. The graph visualizes the data within a certain window length. If data exceeds this window, then the graph is shifted so that the new data can be viewed.

Figure 3.13 Patient-doctor interaction through instantaneous chat.

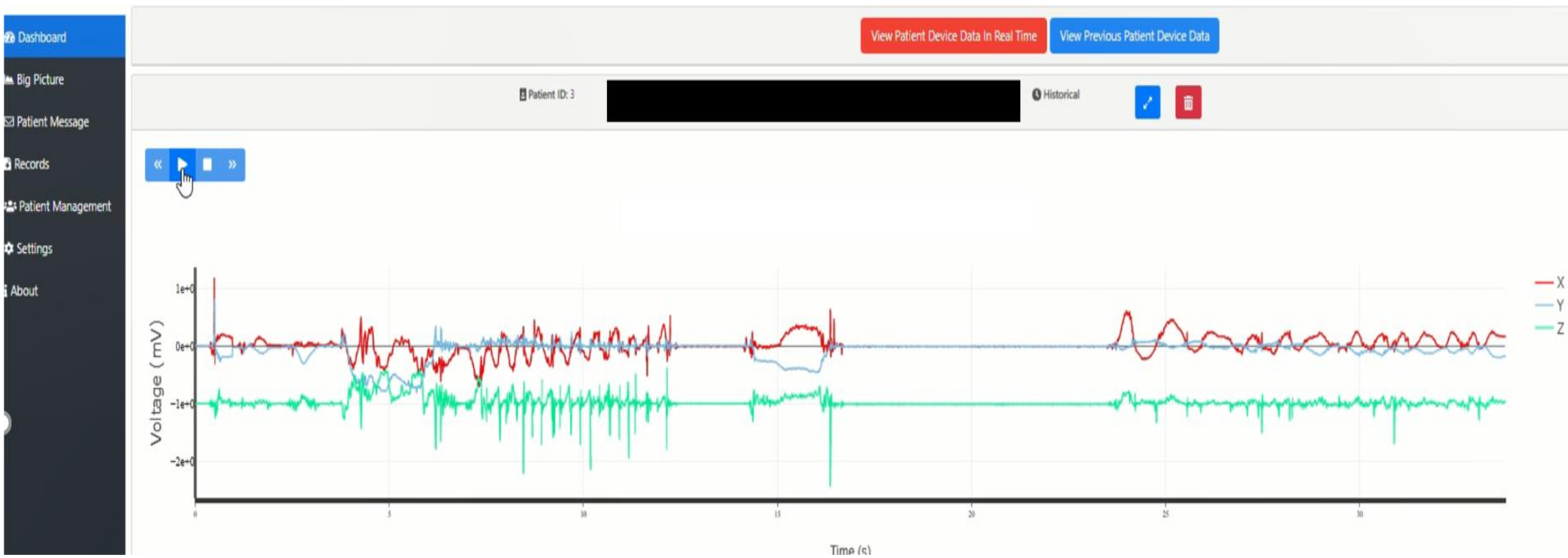

Figure 3.14 Visualization of big data

For the historical case, existing patient data in the cloud storage was used to visualize. The doctor can view data from a specific patient. To view data, the doctor makes selections using the historical visualization form on the Dashboard. Once the information is selected, a request is submitted to retrieve the stored patient data. This request kicks off the data retrieval process. The request is taken in and the data is then retrieved. The data is down-sampled and returned to the user and the website creates graphs to visualize the retrieved data. The data is visualized using a time-series graph for the input data.

The doctor is given a set of buttons that allow them to scroll through the data as shown in Figure 3.14. As the doctor scrolls forward through the data, a set interval of data is retrieved each time. Each time a request is made, the retrieved data is appended to the data that is already stored in the graph. The doctor can also play back the sequence of data that came in from the patient using the play and stop buttons. While playing, the graph will index through the saved data and add the data to the visualization. If it reaches the end of the data, then it will make a request to get more data. Note that the graphs only visualize a certain window of data at a time, so if more data is added then the graph adjusts by shifting the data to the left so that the new data is displayed. When the doctor is done viewing the data of the patient, they can use the delete button to remove the visualization from the Dashboard page. The graphical and data retrieval functions were tested to ensure the doctor could receive, view, and examine the data.

3.4 Conclusion

In this study, we have proposed and developed an innovative smart and secure big data management system and demonstrated its potential on the healthcare big data application. With the security in mind, we have developed functions for doctor sign up and log in, as well as data transmission protection. Further, we have developed the interactive edge-cloud interface, allowing the doctors to conveniently access the big data, and communicate the patients. Besides, both patient information and patient-specific record information can be effectively created and managed. The system further supports both real-time and historical demonstration of the big data, enabling the doctors to effectively leverage the big data. Overall, the proposed s^2Cloud system, empowered by smart and secure design innovations, is expected to advance not only the healthcare big data application, but also many other promising areas.

4. CONCLUSION

Healthcare plays a vital role in everyone's life. This thesis, by innovatively designing the smart and secure big data system, has made two major contributions to the healthcare big data area. First, this thesis has demonstrated the smart and secure mobile edge system. This edge system, by leveraging AWS security features, can provide a secure way for patient registration and management, as well as data protection. Furthermore, the system is highly interactive and allows users (like patients) to examine their data and records, as well as collaborate with cloud users (like physicians). Second, the thesis has demonstrated the smart and secure cloud big data system. With the security in mind, we have developed functions for doctor sign up and log in, as well as data transmission protection. We have developed an interactive edge-cloud interface, allowing doctors to conveniently access the big data, and communicate their patients. Further, effective creation and management of both doctor and patient information is supported by the system. The system can also support both real-time and historical visualization of the big data. Overall, the mobile edge system, and the cloud big data system, are expected to greatly advance the healthcare big data area, and other relevant big data fields.

5. FUTURE STUDIES

In order to boost healthcare systems, researchers must utilize modern technology to create systems that support the growing needs of patients. Despite the number of promising efforts and results this thesis has demonstrated; future work is necessary to further improve the proposed systems. Future studies may include the enhancement of the interactive big data system, by making the system more user friendly and support more functions. The visualization function of the mobile edge and the website can be further improved, to provide more information and more easy-to-use operations. The big data management can be more effective considering the amount of data can be huge and ever increasing. There are other types of data that can be captured on the patient, and each of these types can be used to better analyze patient health. The communication function between the edge and the cloud can also be improved to facilitate more convenient interactions between the patients and doctors. A final future study could be to provide more modalities of big data.

REFERENCES

[1] M. Z. A. Bhuiyan, A. Zaman, T. Wang, G. Wang, H. Tao, and M. M. Hassan, "Blockchain and big data to transform the healthcare," in Proceedings of the International Conference on Data Processing and Applications, 2018, pp. 62-68.

[2] L. Wang and C. A. Alexander, "Big data analytics in medical engineering and healthcare: methods, advances and challenges," J Journal of medical engineering technology and Health Care, vol. 44, no. 6, pp. 267-283, 2020.

[3] M. I. Razzak, M. Imran, and G. Xu, "Big data analytics for preventive medicine," J Neural Computing Applications, vol. 32, no. 9, pp. 4417-4451, 2020.

[4] B. K. Sarkar, "Big data for secure healthcare system: a conceptual design," J Complex Intelligent Systems, vol. 3, no. 2, pp. 133-151, 2017.

[5] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile edge computing: A survey," IEEE Internet of Things Journal, vol. 5, no. 1, pp. 450-465, 2017.

[6] S. Wang, Y. Zhao, J. Xu, J. Yuan, and C.-H. Hsu, "Edge server placement in mobile edge computing," Journal of Parallel and Distributed Computing, vol. 127, pp. 160-168, 2019.

[7] S. Wang, X. Zhang, Y. Zhang, L. Wang, J. Yang, and W. Wang, "A survey on mobile edge networks: Convergence of computing, caching and communications," Ieee Access, vol. 5, pp. 6757-6779, 2017.

[8] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen, and M. Chen, "In-edge ai: Intelligentizing mobile edge computing, caching and communication by federated learning," IEEE Network, vol. 33, no. 5, pp. 156-165, 2019.

[9] J. Huttunen, J. Jauhiainen, L. Lehti, A. Nylund, M. MARTIKAINEN, and O. Lehner, "Big data, cloud computing and data science applications in finance and accounting," ACRN Oxford, JOURNAL OF FINANCE & RISK PERSPECTIVES, vol. 8, 2019.

[10] I. Memon, H. Fazal, R. A. Shaikh, G. Muhammad, Q. A. Arain, and T. K. Khatri, "Big data, Cloud and 5G networks create smart and intelligent world: A survey," University of Sindh Journal of Information and Communication Technology, vol. 3, no. 4, pp. 185-192, 2019.

[11] C. Moreno, R. A. C. González, and E. H. Viedma, "Data and artificial intelligence strategy: A conceptual enterprise big data cloud architecture to enable market-oriented organisations," IJIMAI, vol. 5, no. 6, pp. 7-14, 2019.

[12] S. Rajeswari, K. Suthendran, and K. Rajakumar, "A smart agricultural model by integrating IoT, mobile and cloud-based big data analytics," in 2017 international conference on intelligent computing and control (I2C2), 2017, pp. 1-5: IEEE.

[13] X. Chen, L. Jiao, W. Li, and X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," IEEE/ACM Transactions on Networking, vol. 24, no. 5, pp. 2795-2808, 2015.

[14] S. H. Almotiri, M. A. Khan, and M. A. Alghamdi, "Mobile health (m-health) system in the context of IoT," in 2016 IEEE 4th international conference on future internet of things and cloud workshops (FiCloudW), 2016, pp. 39-42: IEEE.

[15] Q. Zhang, "Mobile Daily Physical Activity and Fall Risk Monitor for Lifestyle Management, Disease Prevention and Risk Alert Generation," Circulation, vol. 141, no. Suppl_1, pp. AP501-AP501, 2020.

[16] J. Stauffer, Q. Zhang, and A. Comer, "Deep Reconstruction Learning Towards Wearable Biomechanical Big Data," in IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES 2021), 2020.

[17] N. H. B. A. Bakar, K. Abdullah, and M. R. Islam, "Wireless smart health monitoring system via mobile phone," in 2016 International Conference on Computer and Communication Engineering (ICCCE), 2016, pp. 213-218: IEEE.

[18] S. D. Nanhore and M. M. Bartere, "Mobile phone sensing system for health monitoring," International Journal of Science and Research, vol. 2, no. 4, pp. 252-255, 2013.

[19] M. S. Mahmud, H. Wang, A. Esfar-E-Alam, and H. Fang, "A wireless health monitoring system using mobile phone accessories," IEEE Internet of Things Journal, vol. 4, no. 6, pp. 2009-2018, 2017.

[20] P.-C. Hii, S.-C. Lee, T.-H. Kwon, and W.-Y. Chung, "Smart phone based patient-centered remote health monitoring application in wireless sensor network," Sensor Letters, vol. 9, no. 2, pp. 791-796, 2011.

[21] N. Oliver, F. Flores-Mangas, and R. De Oliveira, "Towards wearable physiological monitoring on a mobile phone," in Mobile Health Solutions for Biomedical Applications: IGI Global, 2009, pp. 208-243.

[22] AWS. (2021). Amazon Web Service. Available: https://aws.amazon.com/. Last Accessed: 05/26/2021.

[23] i-scoop.eu. (2021). Data Age 2025: the datasphere and data-readiness from edge to core. Available: https://www.i-scoop.eu/big-data-action-value-context/data-age-2025-datasphere/. Last Accessed: 05/26/2021.

[24] X. Zhang et al., "Big data science: opportunities and challenges to address minority health and health disparities in the 21st century," Ethnicity & disease, vol. 27, no. 2, p. 95, 2017.

[25] K. G. Abraham, R. S. Jarmin, B. Moyer, and M. D. Shapiro, "Introduction to" Big Data for 21st Century Economic Statistics"," in Big Data for 21st Century Economic Statistics: University of Chicago Press, 2019.

[26] Z. S. Wong, J. Zhou, and Q. Zhang, "Artificial intelligence for infectious disease big data analytics," Infection, disease & health, vol. 24, no. 1, pp. 44-48, 2019.

[27] A. Alharthi, V. Krotov, and M. Bowman, "Addressing barriers to big data," Business Horizons, vol. 60, no. 3, pp. 285-292, 2017.

[28] Q. Zhang, "Deep Learning-powered Wearable Electrocardiogram Big Data Monitoring for Precision Cardiac Health," Circulation, vol. 141, no. Suppl_1, pp. AP502-AP502, 2020.

[29] Q. Zhang, D. Arney, J. M. Goldman, E. M. Isselbacher, and A. A. Armoundas, "Design Implementation and Evaluation of a Mobile Continuous Blood Oxygen Saturation Monitoring System," Sensors, vol. 20, no. 22, p. 6581, 2020.

[30] J. Zou and Q. Zhang, "Intelligent Mobile Electrocardiogram Monitor-empowered Personalized Cardiac Big Data," in The 11th IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (IEEE UEMCON), 2020.

[31] J. Liu, F. Xie, Y. Zhou, Q. Zou, and J. Wu, "A wearable health monitoring system with multi-parameters," in 2013 6th International Conference on Biomedical Engineering and Informatics, 2013, pp. 332-336: IEEE.

[32] T. Suzuki, H. Tanaka, S. Minami, H. Yamada, and T. Miyata, "Wearable wireless vital monitoring technology for smart health care," in 2013 7th International Symposium on Medical Information and Communication Technology (ISMICT), 2013, pp. 1-4: IEEE.

[33] L. Zhang, L. Yang, Z. Wang, and D. Yan, "Sports wearable device design and health data monitoring based on wireless internet of things," Microprocessors and Microsystems, p. 103423, 2020.

[34] H. Yan, Y. Jiang, J. Zheng, C. Peng, and Q. Li, "A multilayer perceptron-based medical decision support system for heart disease diagnosis," Expert Systems with Applications, vol. 30, no. 2, pp. 272-281, 2006.

[35] M. Fieschi, J.-C. Dufour, P. Staccini, J. Gouvernet, and O. Bouhaddou, "Medical decision support systems: old dilemmas and new paradigms?," Methods of information in medicine, vol. 42, no. 03, pp. 190-198, 2003.

[36] B. Malmir, M. Amini, and S. I. Chang, "A medical decision support system for disease diagnosis under uncertainty," Expert Systems with Applications, vol. 88, pp. 95-108, 2017.

[37] M. C. Roziqin, D. S. H. Putra, and M. S. Noor, "Information System for Doctor Practice Scheduling at Hospitals in Jember District," in The First International Conference on Social Science, Humanity, and Public Health (ICOSHIP 2020), 2021, pp. 29-31: Atlantis Press.

[38] V. Vysotska, V. Lytvyn, Y. Burov, A. Gozhyj, and S. Makara, "The Consolidated Information Web-Resource about Pharmacy Networks in City," in IDDM, 2018, pp. 239-255.

[39] Y. Inal, "Heuristic-based user interface evaluation of the mobile centralized doctor appointment system," The Electronic Library, 2019.

[40] A. Nursikuwagus, "E-Health as a Service Software of Medical System in Prototype Modeling," IJNMT (International Journal of New Media Technology), vol. 4, no. 2, pp. 99-104, 2017.

[41] Amazon. (2018). Amazon Web Service. Available: https://aws.amazon.com/. Last Accessed: 05/26/2021.

[42] G. Manogaran, C. Thota, D. Lopez, and R. Sundarasekar, "Big data security intelligence for healthcare industry 4.0," in Cybersecurity for Industry 4.0: Springer, 2017, pp. 103-126.

[43] K. Abouelmehdi, A. Beni-Hssane, H. Khaloufi, and M. Saadi, "Big data security and privacy in healthcare: a review," Procedia Computer Science, vol. 113, pp. 73-80, 2017.

[44] K. R. Sollins, "IoT big data security and privacy versus innovation," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 1628-1635, 2019.

[45] C. Thota, G. Manogaran, D. Lopez, and V. Vijayakumar, "Big data security framework for distributed cloud data centers," in Cybersecurity breaches and issues surrounding online threat protection: IGI global, 2017, pp. 288-310.

[46] H. Mahmoud, A. Hegazy, and M. H. Khafagy, "An approach for big data security based on Hadoop distributed file system," in 2018 International Conference on Innovative Trends in Computer Engineering (ITCE), 2018, pp. 109-114: IEEE