



Cloud-Based Digital Twins' Storage in Emergency Healthcare

Erdan Wang¹ · Pouria Tayebi¹ · Yeong-Tae Song¹

Received: 28 June 2023 / Accepted: 8 August 2023 / Published online: 27 August 2023
© The Author(s) 2023

Abstract

In a medical emergency situation, real-time patient data sharing may improve the survivability of a patient. In this paper, we explore how Digital Twin (DT) technology can be used for real-time data storage and processing in emergency healthcare. We investigated various enabling technologies, including cloud platforms, data transmission formats, and storage file formats, to develop a feasible DT storage solution for emergency healthcare. Through our analysis, we found Amazon AWS to be the most suitable cloud platform due to its sophisticated real-time data processing and analytical tools. Additionally, we determine that the MQTT protocol is suitable for real-time medical data transmission, and FHIR is the most appropriate medical file storage format for emergency healthcare situations. We propose a cloud-based DT storage solution, in which real-time medical data are transmitted to AWS IoT Core, processed by Kinesis Data Analytics, and stored securely in AWS HealthLake. Despite the feasibility of the proposed solution, challenges such as insufficient access control, lack of encryption, and vendor conformity must be addressed for successful practical implementation. Future work may involve Hyperledger Fabric technology and HTTPS protocol to enhance security, while the maturation of DT technology is expected to resolve vendor conformity issues. By addressing these challenges, our proposed DT storage solution has the potential to improve data accessibility and decision-making in emergency healthcare settings.

Keywords Digital twins · Cloud computing · Storage · Real-time data stream · Emergency healthcare · FHIR

1 Introduction

In contemporary emergency healthcare, there is a scarcity of effective methods for sharing patient information between physicians in the Emergency Department (ED) and paramedics in an ambulance. Consequently, patient demographics, chief complaints, vital signs, treatment responses, and other medical data are typically transmitted through verbal or written reports. This mode of communication could potentially result in miscommunication or the omission of crucial medical information, particularly in time-sensitive situations such as myocardial infarctions. Traditional Electronic Health Record (EHR) and Electronic Medical Record

(EMR) systems can store and share static medical records but lack the capability to process and share real-time medical data, such as electrocardiograms (EKG) and peripheral capillary oxygen saturation (SpO₂). Employing patients' Digital Twins (DT) could address this challenge, as DT can store, process, analyze, and share data in real time. Moreover, paramedics might leverage patients' DTs to simulate treatment and medication responses.

Although the definition of digital twins varies according to different authors' perspectives, the fundamental concepts remain consistent. Digital twins represent virtual replicas of physical objects that facilitate sharing, analyzing, optimizing, and predicting the performance of their real-world counterparts [1]. Fundamentally, digital twins comprise the digital model of a physical object, historical data, and real-time data. Real-time data processing (unidirectional or bidirectional) is a critical element that differentiates DT from other digital approaches, such as simulation and digital modeling. DT enables bidirectional human-out-of-the-loop communication between the physical system and the virtual model. Data transfer from a physical object to the DT is processed, analyzed, and evaluated, potentially

✉ Erdan Wang
ewang6@students.towson.edu
Pouria Tayebi
ptayebi1@students.towson.edu
Yeong-Tae Song
ysong@towson.edu

¹ Department of Computer and Information Science, Towson University, Towson, USA

offering predictions for the object's performance and status. Additionally, data feedback from the DT to the physical object can facilitate notifications or control actions. Within an emergency healthcare context, DT could be employed to store and share a patient's real-time vital sign data from the ambulance and static medical records, including demographics, chief complaints, treatments, and medications. Furthermore, DT offers potential simulation and estimation capabilities, which could be pivotal in determining patient treatment and medication choices.

Before embarking on any medical treatment or therapy, one key aspect of healthcare involves understanding how the patient's body will react to the proposed intervention. This can entail evaluating the patient's current health status, taking note of any pre-existing conditions or allergies, and considering their medical history. This initial analysis is paramount as it can help to predict potential side-effects and adverse reactions to the therapy, especially those that could worsen the patient's condition or impair their recovery.

With thorough understanding of the patient's physiological condition and potential reaction with the proposed treatment, healthcare providers can make more informed decisions. They can tailor the therapy to the patient's specific needs and circumstances, potentially altering dosages, delivery methods, or even opting for different treatments entirely based on their analysis. In turn, this individualized approach to treatment can improve patient outcomes, reduce the likelihood of adverse reactions, and increase overall patient satisfaction with their care.

In emergency healthcare, digital twins can be used to share real-time condition for patient's vital signs or individual body systems such as cardiovascular system with other emergency caregivers. When integrated with other patient data—such as previous medical history, genetic information, lifestyle factors, and physiological condition—these digital replicas can provide invaluable insights for healthcare providers.

Digital twins of a patient can be created at many levels. A digital twin at the level of the entire human body would be a comprehensive, integrated model encompassing all of the body's systems and organs. This would ideally involve extensive data collection to capture as much detail as possible about the individual's health, including medical history, genetic data, lifestyle factors, and current physiological status. A full-body digital twin could be used to simulate the effects of a wide range of treatments and interventions, allowing clinicians to predict and prevent potential complications or adverse reactions. Furthermore, it could be utilized for monitoring overall health and predicting potential health risks, contributing to preventative care.

Creating digital twins at the level of body functions or systems, such as the respiratory, nervous, or urinary systems, allows for the detailed analysis of how these systems

function and interact with one another. For instance, a digital twin of the respiratory system could be used to study the progression of diseases like COPD or asthma and to predict how the system might respond to various treatments. Similarly, a digital twin of the nervous system could be used to study neurological disorders, and a digital twin of the urinary system could help in understanding and managing conditions like kidney disease.

Creating digital twins at the organ level involves modeling individual organs such as the heart, kidney, or liver. This can be especially useful for predicting the outcomes of organ-specific treatments, including surgeries, medication regimens, or innovative procedures like organ transplantation. For example, a digital twin of a patient's heart could help physicians better understand how the heart might respond to different types of cardiac surgeries or medications, helping to optimize patient outcomes.

At the lowest level, digital twins can be created to represent the smallest components of the body, such as cells and molecules. This can be particularly useful in the field of oncology, where understanding the behavior of cancer cells and their interaction with various drugs is crucial. Similarly, a digital twin at the molecular level could be used to study the effects of genetic variations on disease development and response to treatment. It could also assist in drug development by simulating how different molecules interact with each other and with biological targets.

Creating digital twins at various levels of the human body allows for a highly personalized and detailed understanding of individual health. By providing precise, individualized models of different parts of the body, they offer significant potential for enhancing the accuracy of diagnoses, the effectiveness of treatments, and the overall quality of patient care. This reflects the broader move in healthcare toward precision medicine, which aims to tailor medical treatment to the individual characteristics of each patient.

A DT model can also belong to one of two categories: Healthy Baseline Models or Abnormal Disease Models. A healthy baseline digital twin model represents the ideal or normal state of health for a specific individual. This model is created based on comprehensive data collected from the individual when they are in a state of good health, including genetic data, physiological data, lifestyle factors, and other relevant health indicators. The healthy baseline model serves as a benchmark or standard against which future health status can be compared.

Regular updates to the model ensure that it evolves as the individual's health changes over time, thus maintaining its accuracy and relevance. The main purpose of this model is to understand what constitutes 'normal' for that individual and identify any deviations from this norm. This can help detect potential health issues at an early stage, allowing for proactive interventions and preventative care. Furthermore,

the model can be used to simulate how the individual's body would react to different treatments or lifestyle changes, aiding in personalized healthcare.

An abnormal disease digital twin model represents the state of health of an individual who is affected by a specific disease or medical condition. This model is created based on extensive data collected from the individual, including the nature and severity of their illness, their medical history, genetic data, and current physiological status. The abnormal disease model provides a detailed picture of the disease process within the individual's body, helping clinicians to understand how the disease is likely to progress and how the body may respond to different treatments.

This type of model can be particularly useful for complex or chronic diseases where treatment needs to be highly personalized, such as cancer or heart disease. It allows for the simulation of different treatment options and their likely outcomes, enabling more informed decision-making and potentially improving patient outcomes. In addition, the model can be used to monitor the effectiveness of treatment over time and adjust the treatment plan as necessary.

Both healthy baseline models and abnormal disease models are critical in modern healthcare. They enable a highly personalized approach to care, facilitating early detection of health issues, informed treatment decision-making, and effective monitoring of treatment progress. Digital twins represent an exciting development in the field of precision medicine, with the potential to significantly enhance the quality and effectiveness of healthcare.

One of the primary benefits of using digital twins in healthcare is their potential to increase efficiency and precision in patient treatment, thereby reducing the incidence of medical errors. Through simulations, clinicians can explore different treatment options and predict their outcomes before applying them to the actual patient. For instance, a digital twin can help simulate the effects of a new drug on a specific patient, or how a surgical intervention might impact the patient's body. This reduces the risk of unforeseen complications or adverse reactions, as potential issues can be addressed before the actual therapy or procedure is implemented.

Moreover, digital twins can facilitate more proactive and preventative care by predicting potential health issues before they become severe. For example, by monitoring changes in the digital twin over time, healthcare providers may be able to detect the early stages of a disease that has yet to manifest noticeable symptoms in the patient. As a result, treatment can commence sooner, often leading to better outcomes.

In addition, the use of digital twins can enhance the efficiency of healthcare delivery. By providing a clear and comprehensive overview of a patient's health status, they enable healthcare providers to make quicker, more accurate diagnoses, devise more effective treatment plans, and

monitor patient progress more closely. This not only leads to improved patient care but also optimizes the use of healthcare resources.

Overall, the integration of digital twins in healthcare represents a significant stride toward more personalized, efficient, and accurate care, underpinning the move toward precision medicine. Their use can help to reduce medical errors, improve patient outcomes, and enhance the overall efficiency of the healthcare system.

This paper presents a cloud-based DT solution to address the challenges of storing patients' medical information (both real-time and static) in emergency healthcare settings. The subsequent sections of the paper encompass a review of related works, an examination of the enabling technologies, a discussion of the proposed solution, and a conclusion.

2 Related Works

2.1 Digital Twins and Healthcare

Although Digital Twins (DT) represent a relatively nascent technology, researchers have already investigated its application within the healthcare domain, yielding promising outcomes. Song et al. [2] conducted a survey on the utilization of DT in personal healthcare, proposing its use for assessing and maintaining individual health. They argued that the quality of a DT is contingent upon the completeness, update frequency, and accuracy of the observed data, which encompasses personalized risk factors such as behavior, vital signs, or symptoms. Xie et al. [3] applied DT and data mining techniques to electroencephalogram (EEG) data and medical history from the SEER database of patients, subsequently providing health status and feedback. Rivera et al. [4] outlined their vision for implementing DT in precision medicine, developing a reference model that capitalizes on DT capabilities, self-adaptive systems, and autonomic computing to engineer intelligent and flexible software systems for healthcare. The authors also delineated internal structures for DT to support precision medicine methodologies in the context of continuous monitoring and personalized, data-driven medical treatments.

2.2 Digital Twins and Healthcare Existing Problems in Communication Between Emergency Medical Services (EMS) and Emergency Department (ED)

The communication systems deployed between EMS and ED currently are verbal over radio and written through designated forms. Studies suggest that the handover of information can be diluted due to the stressful environment the paramedics are communicating with the ED. Additionally,

the progress and/or changes of patient's appearance from the very first instance observed by paramedics to the point of hospital admission must be recorded, which is occasionally disregarded by current communication system. A primary solution proposed by Troyer et al. [5] is integrated medical record shared between medical center and EMS, which has not been studied and does not eliminate real-time data transfer, whilst some of the problems such as the effect of environment on data recording will remain in place.

EMS utilizes two channels of data communication for handing over the patient to ED. Firstly, Transfer of Care (TOC) form is filed when transferring the patient to ED. The latter is a Patient Care Report (PCR), which is filed up to 72 h after care is concluded. On average, there is 25% discrepancies (such as allergies, used medications, etc.) between the two reports. In specific, numerical data such as heart rate and blood pressure were mismatched between TOC and PCR, and major data losses in both reports, with TOC missing around 40% more data than PCR [6]. This is concerning as TOC should be as precise as possible for urgent medical care. Improving data accuracy and assuring its completeness at the point of transferring patients to ED can be clearly seen to be an improving factor in providing faster medical service with higher quality inside the hospital.

Additionally, EMS data recording and its communication with ED faces numerous more challenges such as difficulty of obtaining an unconscious patients' information, paramedics' hardship in multi-tasking for providing care and passing information to hospital, radio signal instability, and communication overload at peak hours [7]. A solution proposed in this paper, adopting technology to overcome these challenges, is developing a mobile app to communicate with ED through captured videos (this was suggested to be through smart glasses) and photographs of patients alongside voice notes of paramedics. However, the mobile app will not be sending real-time data and their synchronization with any given medication to the patient.

2.3 Use of Digital Twins (DTs)

The novel concept of digital twins has been considered/ deployed in different applications ranging from industrial manufacturing [8] to ED [9] and human body [10]. The level of DT's sophistication varies through each application due to the incompleteness of real-time and detailed data obtainable from Physical Twin (PT) in certain applications such as the human body as a whole. With our focus being on human body DT development, data and measurements associated with different body organs must be collected, this can be through real-time sensor data, medical history, observations, and genetic information [10]. An example of developed DT model on human heart using continuous obtainable data (ECG, blood pressure and heart rate) alongside uniform data

such as heart geometry obtained through MRI, has been proposed by Kayvanpour et al. [11].

Sahal et al. [10] propose a framework to combine DT, AI and blockchain to improve physician–patient communication, facilitate shared decision-making and data privacy protection, respectively. A range of probable benefits in developing human body DT ranging from digital patient model, personalized treatment, rapid diagnosis, predicting response to surgical interventions, joint research opportunities and ultimately, human immortality through data were discussed. However, this paper can be seen as a general guidance and doesn't attempt to solve a specific problem in healthcare, such as EMS patient flow, or methods of DT creation and their data storage.

2.4 Digital Twin Data Structure

Singh et al. [12] proposed method of data structure that utilizes ontology for relational database management and transformation. They specifically use ontology to database mapping to overcome issues related to continuous update of data, easing the data query on the user's side. Ontology to database mapping assumes the pre-existence of database and ontology, which produces a set corresponding mapping between the two schemas. Additionally, they have chosen relational database over graph-based data representation to support multi-user environment and standardized query language, whilst the graph-based approach has been admitted being more suitable for dynamic data modeling that are highly connected and complex, which is the case for DT data representation. Digital twin technology explored in healthcare, has provided little findings on practical approach for data structure [13] but detailed framework in how to store big data in different layers in cloud has been found in literature [14].

2.5 Digital Twin Data Storage and Transmission

Digital twin data storage varies based on its application. Glancing through industrial and healthcare applications, if data storage topic was touched on, cloud-based models were proposed [8, 15]. Although, there are different cloud platforms that can be used, and the suitability varies based on the application's requirements. In developing a DT of human heart, Edge storage and computing platform was proposed [15] to reduce latency to a minimum in data analytics. The significance of reduced latency would be highlighted in urgent scenarios.

Other forms of data storage such as distributed file storage (DFS), NoSQL and NewSQL databases are proposed in literature [16], which impose little emphasis on data security and focus on available methodologies that can be used in different applications. In the same paper, the writers

propose Aspera tool for data transmission, with underlying secure communication protocols such as HTTPS, SSL or SFTP. Hence, it is thought to be more suitable for real-time data flow. Other similar communication technologies such as LoRaWAN, 6G have been mentioned for data communication over the internet [17], whilst the implementation of lower layer protocols such as MQTT and ensuring its security for DT has not been found in literature.

2.6 Internet of Things and Cloud Computing

Various technologies can facilitate DT implementation, including the Internet of Things (IoT), cloud computing platforms, edge computing, APIs and integration platforms, machine learning and AI, and blockchain and distributed ledger technologies. This paper specifically concentrates on IoT and cloud computing as enablers of DT storage. Khaled et al. [18] designed a motor and wind turbine DT model using Simscape and Simulink, subsequently deploying the DT model on the Amazon AWS cloud. Saad et al. [19] proposed a cloud-based DT solution for power grid distribution systems, importing DT models into the AWS Lambda function and employing AWS Greengrass at the edge to extend cloud system functionality to IoT devices. Wang et al. [20] suggested an AWS-based Mobile Digital Twin (MDT) solution for personalized adaptive cruise control. The MDT (Fig. 1), comprising multiple DTs and interactions among them, represents a relatively comprehensive IoT and cloud solution for DT. The concept of cloud-based DT was validated and tested, with results demonstrating the feasibility of

real-time DT implementation based on IoT and cloud computing technologies.

3 Enabling Technologies

The focus of this paper is on identifying a practical and viable solution for storing Digital Twins (DT) in emergency healthcare. To this end, several technologies pertinent to DT cloud storage are surveyed, including cloud platforms, data transmission formats, and file storage formats.

3.1 Cloud Platforms

Table 1 presents the results of the cloud platform survey. The table juxtaposes DT storage-related technologies from three major cloud service providers. In terms of services for medical file storage and processing, Amazon offers Healthcare Lake services, Microsoft Azure features Azure API for FHIR, and Google Cloud provides the Google Cloud Healthcare API. These services facilitate the storage, management, and analysis of health data in the cloud while adhering to standards such as FHIR, DICOM, and HL7.

Regarding DT services, Amazon introduces AWS IoT TwinMaker, Microsoft supplies Azure Digital Twins and Digital Twin Definition Language, and Google offers Supply Chain Twin. These services enable the creation, visualization, and management of digital twins for IoT devices and systems, allowing for real-time synchronization,

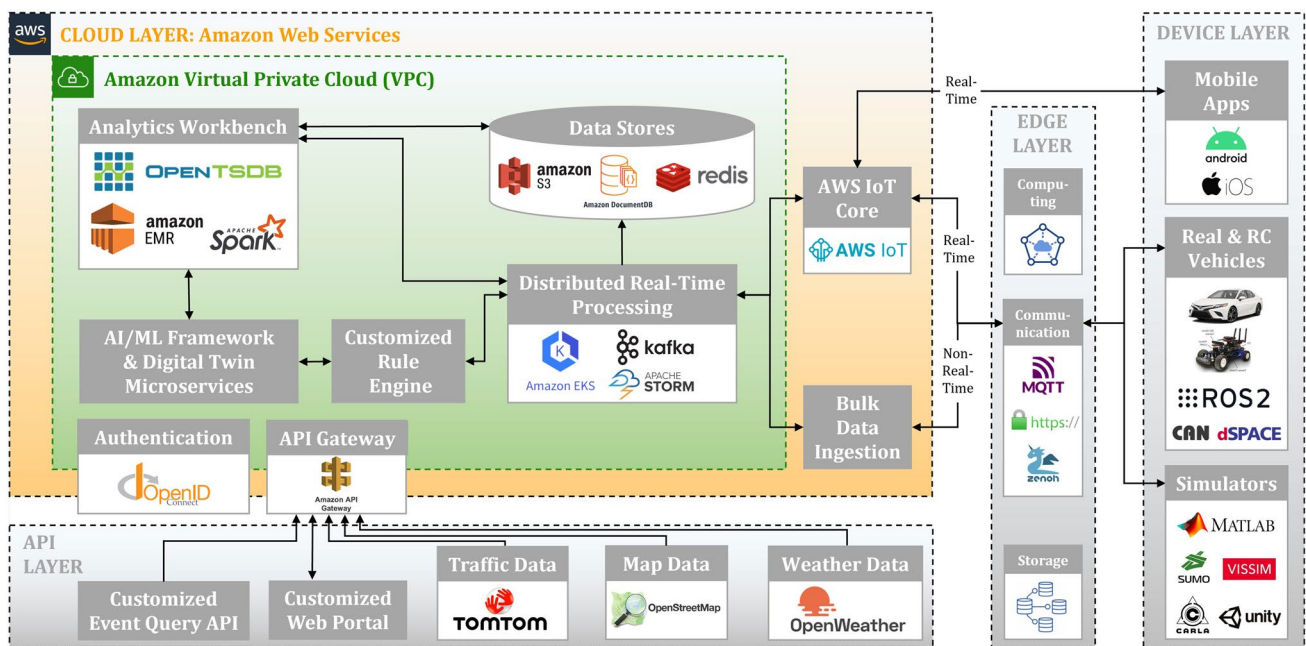


Fig. 1 MDT cloud solution proposed by Wang et al. [20]

Table 1 The survey of cloud services

Cloud provider	Surveyed services		
	Medical files storage	Real-time data stream processing	DT
AWS	HealthcareLake	Amazon kinesis data Stream/analytics	Amazon IoT TwinMaker
Azure	Azure API for FHIR	Azure stream analytics	Azure digital twins, Digital twin definition language
Google Cloud	Google Cloud Healthcare API	Cloud streaming analytics	Supply chain twin
Cloud provider	Surveyed services		
	IoMT	Data analysis	HIPAA
AWS	No specific IoMT services, but can use IoT services	Amazon Athena, Amazon Quick Sight	Yes ^a
Azure	No Specific IoMT services, but can use IoT services ^d	Azure Machine learning	Yes ^b
Google Cloud	No Specific IoMT services, but can use IoT services	Cloud AutoML	Yes ^c

^aAWS has a standard Business Associate Addendum (BAA)

^bAzure has a standard Business Associate Addendum (BAA)

^cGoogle Cloud has a standard Business Associate Addendum (BAA)

^dAzure has a GitHub IoMT FHIR connector

enhanced analytics, predictive maintenance, and improved decision-making.

Real-time data stream processing is a vital component for storing and processing real-time data generated by medical devices in emergency healthcare settings. Amazon provides Kinesis data streams/analytics, Microsoft offers Azure Stream Analytics, and Google features Cloud Streaming Analytics. These services ensure efficient and reliable data processing and analysis, empowering data-driven decision-making and rapid response to changing conditions.

In emergency healthcare environments, all data transmission and storage must comply with HIPAA regulations. This necessitates the encryption and validation of data generated by medical devices. While none of the surveyed cloud platforms offer specific IoMT services, they all assert that their IoT services comply with HIPAA. Each cloud platform has signed a standard Business Associate Addendum (BAA), implying that the services listed under the BAA adhere to HIPAA regulations.

3.2 Data Transmission Format

In emergency healthcare settings, the Message Queuing Telemetry Transport (MQTT) protocol ensures real-time transmission of data streams generated by medical equipment. MQTT is a lightweight, publish-subscribe messaging protocol designed for efficient and reliable communication between IoT devices, even in constrained or low-bandwidth environments.

3.3 Storage File Format

Upon arriving at the cloud platforms, the real-time medical data stream is subjected to various processing steps and directed to two distinct destinations. The first destination entails immediate real-time processing and transmission to the Emergency Department (ED) of the recipient hospital. Conversely, the second destination involves the storage of the data within the Digital Twin (DT) residing in the cloud. The real-time medical data may be stored in one of two file formats within the DT.

3.3.1 HL7 C-CDA

Health Level Seven Consolidated Clinical Document Architecture (HL7 C-CDA) is a standard for the exchange of electronic health information in a structured, interoperable format. It builds upon the Clinical Document Architecture (CDA) standard, which defines the structure and semantics of clinical documents, such as discharge summaries, progress notes, and history and physical reports.

C-CDA is a part of the broader HL7 framework, a set of international standards for the exchange, integration, sharing, and retrieval of electronic health information. HL7 C-CDA provides a unified approach to structuring clinical documents by consolidating multiple CDA implementation guides into a single, harmonized standard.

3.3.2 HL7 FHIR

HL7 FHIR, an acronym for Health Level Seven Fast Healthcare Interoperability Resources, embodies a contemporary, web-centric standard dedicated to facilitating the exchange of healthcare information between disparate systems. It is characterized by its emphasis on secure, efficient, and interoperable data transmission. Its development is anchored in the recognition of the paramount importance of data sharing in the modern healthcare landscape, particularly given the proliferation of digital health records and advancements in data-driven care.

The FHIR standard is engineered to confront and navigate the myriad complexities and challenges intrinsic to the field of healthcare data exchange. The idiosyncrasies of healthcare data—its diversity, sensitivity, and the imperative for accuracy—make its handling a uniquely challenging endeavor. FHIR, as a standard, is designed to address these very challenges, bringing about a simplified, standardized, and interoperable approach to healthcare data exchange.

One of the core features of FHIR is its resource-oriented architecture, which encapsulates discrete healthcare concepts into individual 'resources'. Each resource contains a specific, well-defined chunk of healthcare data, such as a patient's demographic information, medical history, or laboratory results. This modular design enables the individual resources to be combined and used in various contexts, fostering a high degree of interoperability.

Moreover, FHIR is built with the modern web in mind. It adopts many of the design principles and technologies commonly used in contemporary web services, such as RESTful APIs, JSON, and XML. This design choice not only ensures compatibility with current technologies but also provides developers with familiar tools and methodologies, facilitating the adoption of the standard.

An additional strength of FHIR is its flexibility. It supports a wide range of use cases, from simple data retrieval to complex clinical decision support, making it suitable for a variety of healthcare scenarios. Furthermore, its extensible design allows for the incorporation of new data types or modifications to existing ones, ensuring the standard can evolve alongside advancements in healthcare and technology.

The importance of security in healthcare data cannot be overstated, and FHIR addresses this need head-on. The standard includes built-in provisions for data security and privacy, supporting a range of security models and authorization protocols. It ensures that patient data can be securely exchanged between systems while adhering to relevant privacy regulations and standards.

Through FHIR, healthcare organizations can share and integrate patient data more easily, making it a pivotal enabler of interoperability in healthcare. It facilitates the fluid

exchange of information, thereby fostering a more connected and cohesive healthcare ecosystem. This capability is crucial in the era of digital health, where timely access to comprehensive and accurate patient data can significantly enhance clinical decision-making and, ultimately, patient outcomes.

4 Discussion on Our Proposal

Upon examining the enabling technologies discussed in the preceding section, several observations can be made. All three cloud platforms demonstrate their suitability for constructing DT storage in emergency healthcare. However, since our study primarily focuses on the storage of real-time data within DT, Amazon AWS offers more sophisticated and powerful real-time data processing and analytical tools compared to Microsoft Azure and Google Cloud. Consequently, we have selected AWS as the cloud platform upon which we will construct our DT storage solution. AWS was selected as a platform to create temporal digital twin also due to several other factors explained below.

4.1 Zone Availability

Due to higher market share of AWS compared to its competitors (Microsoft Azure and Google Cloud), has led this service to have greater availability in different zones [21]. Holding a prosperous insight in this project, which means its expansion to different geographical locations, would require a service with highest availability in different zones. Hence, AWS is advantageous in this aspect.

4.2 Security

HIPAA compliancy for data transmission between different modules of AWS is/can be implemented. For instance, certain AWS services such as CloudTrail and Amazon CloudWatch can be used to set logging requirements to ensure sensitive data sent through API Gateway are captured by HIPAA-eligible services. Additionally, storage modules (such as S3 and Glacier) are encrypted using management keys to ensure data privacy [22].

Microsoft Azure can provide layered security architecture for different application tiers to restrict public access to API applications as well as the existence of role-based access control and data encryption for transit and storage [23], which overall allows HIPAA compliancy.

Similar security measures are implementable on Google Cloud [24], showcasing a competing scenario between cloud service providers from security perspective. Although, AWS has a higher reputation in security for addressing exact industrial needs such as HIPAA [22, 25].

4.3 Cost

The cost comparison for strongest instances that AWS and Microsoft Azure have on offer are analyzed and concluded that AWS's processing engine is a cheaper service provider than Azure and Google Cloud [25–27], whilst AWS's IoT platform is more expensive than its competitors due to greater range of cloud services it provides [28].

The MQTT protocol is deemed appropriate for transmitting real-time medical data, and as for the medical file storage format, FHIR is better suited for emergency healthcare scenarios due to its granularity, flexibility, and efficient API, which makes it an optimal choice for real-time data storage and exchange in most situations.

Figure 2 provides a detailed visual representation of the proposed cloud-based Digital Twin (DT) storage solution, specifically designed to cater to the needs of emergency healthcare. This system incorporates an efficient, end-to-end process of data capture, processing, transmission, and storage. The utilization of the Amazon Web Services (AWS) infrastructure, which is known for its robustness and scalability, forms the backbone of this solution.

In this sophisticated system, the initial data are captured by a diverse array of medical equipment aboard an emergency vehicle, such as an ambulance. The information gathered from these medical devices ranges from basic vital signs to more complex data, all of which contribute to painting a comprehensive picture of the patient's health status.

This collected data are subsequently transmitted to a Raspberry Pi device on the ambulance. The Raspberry Pi, a small but versatile computer, performs the function of transforming the captured data into Message Queuing Telemetry Transport (MQTT) messages (Fig. 3). MQTT is a lightweight messaging protocol designed for low-bandwidth, high-latency networks, making it particularly suitable for transmitting data from medical devices in emergency settings.

Following this transformation, the MQTT messages are relayed to the IoT Gateway of AWS IoT Core (Fig. 4) located within the larger AWS Cloud infrastructure. The IoT Gateway serves as the entry point for these messages

into the cloud, effectively bridging the gap between the data source (the ambulance) and the cloud storage and processing system.

The IoT gateway, based on the rules established by IoT Act Rules, then distributes the incoming messages to the Simple Notification Service (SNS) Topic and AWS Kinesis (Fig. 5). The former is a web service that coordinates the delivery or sending of messages to subscribing endpoints, while the latter is a real-time data streaming service offered by AWS.

Upon receiving the data, AWS Kinesis generates a tabular representation of the information, creating a structured format that is amenable to further analysis. This data can then be scrutinized through Kinesis Data Analytics (Fig. 6), a feature that provides insights into the processed data.

Simultaneously, a web server deployed on Amazon Elastic Compute Cloud (Amazon EC2) receives the data stream from Kinesis Data Stream. This information is displayed on a web portal that can be accessed by handheld devices in the Emergency Department (ED) of the recipient hospital, allowing healthcare professionals to view real-time patient medical data.

Additionally, the SNS messages are routed to an AWS Lambda function, a serverless compute service that runs code in response to events and automatically manages the compute resources (Fig. 7). This function transforms the data stream into Fast Healthcare Interoperability Resources (FHIR) files, a standard for exchanging healthcare information electronically.

The resultant FHIR files (Fig. 8) are then stored in an Amazon Simple Storage Service (S3) bucket, providing secure and scalable object storage. The stored data can be subsequently queried using AWS HealthLake, a HIPAA-eligible service for healthcare providers, health insurance companies, and pharmaceutical companies to store, transform, query, and analyze health data at scale. The insights derived from these queries are stored back in the S3 bucket, providing a comprehensive data record for future reference and ongoing patient care.

In conclusion, the proposed cloud-based DT storage solution offers an integrated, efficient, and robust system for

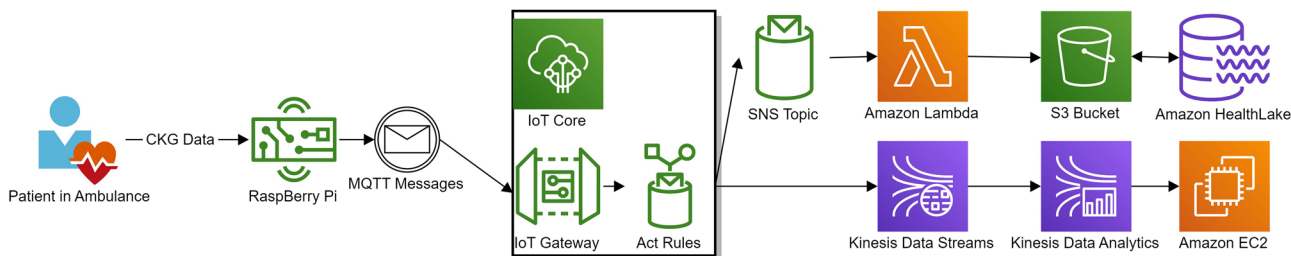


Fig. 2 Proposed cloud-based DT storage solution


```

main.py
25 myAWSIoTClient.subscribe(topic, 1, on_message_received)
26
27
28 # Publish a test message
29 test_message = {
30     "message": "Hello, AWS IoT!"
31 }
32 myAWSIoTClient.publish(topic, json.dumps(test_message), 1)
33
34 # Keep the program running and receiving messages
35 try:
36     while True:
37         time.sleep(1)
38     except KeyboardInterrupt:
39         pass
40
41 # Disconnect the client before exiting
42 myAWSIoTClient.unsubscribe(topic)
43 myAWSIoTClient.disconnect()
    
```

```

Received message: b '{"deviceid": "rpi-700d", "bpm": 80, "spo2": 90, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "John_Doe#1115555", "sequence": 12, "timestamp": 1693015586}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 87, "spo2": 93, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278145941}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 93, "spo2": 92, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278146178}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 77, "spo2": 92, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278146420}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 68, "spo2": 85, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278146833}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 68, "spo2": 89, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278146921}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 73, "spo2": 95, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278147155}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 77, "spo2": 98, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278147465}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 68, "spo2": 85, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278147641}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 62, "spo2": 94, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278147894}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 68, "spo2": 91, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278148132}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 70, "spo2": 98, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278148381}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 77, "spo2": 93, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278148629}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 70, "spo2": 93, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278148877}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 69, "spo2": 95, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278149129}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 85, "spo2": 86, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278149360}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 78, "spo2": 93, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278149613}' from topic: r
Received message: b '{"deviceid": "rpi-700d", "bpm": 68, "spo2": 92, "temperature": 80, "humidity": 30, "latitude": 45.001, "longitude": 1.002, "msg": "name#1115555", "sequence": 12, "timestamp": 1687278149861}' from topic: r
    
```

Fig. 3 The MQTT messages received by a test program that subscribes to our AWS IoT Topic

```

AWS IoT
Monitor
Connect
  Connect one device
  Connect many devices
Test
  MQTT test client
Manage
  All devices
  Greengrass devices
  Software packages
  Remote actions
  Message routing
  Retained messages
  Security
  Fleet Hub
Device software
Billing groups
Settings
Feature spotlight
    
```

```

rpi-healthstation-topic August 25, 2023, 21:59:15 (UTC-0400)
{
  "deviceid": "MAX-30102",
  "bpm": 78,
  "spo2": 96,
  "temperature": 80,
  "humidity": 30,
  "latitude": 45.001,
  "longitude": 1.002,
  "msg": "John_Doe#1115555",
  "sequence": 12,
  "timestamp": 16930155886
}
    
```

```

rpi-healthstation-topic August 25, 2023, 21:59:15 (UTC-0400)
{
  "deviceid": "MAX-30102",
  "bpm": 96,
  "spo2": 92,
  "temperature": 80,
  "humidity": 30,
  "latitude": 45.001,
  "longitude": 1.002,
  "msg": "John_Doe#1115555",
  "sequence": 12,
  "timestamp": 16930155637
}
    
```

Fig. 4 The MQTT messages received by AWS IoT Core

managing critical patient data in emergency healthcare settings. By harnessing the power of cutting-edge technologies and the AWS infrastructure, this system can significantly enhance real-time data availability, processing, and analysis, potentially improving patient outcomes and the overall efficiency of emergency medical services.

Although the proposed solution appears viable, several challenges must be addressed before practical implementation can occur:

- Inadequate access control may allow unauthorized and potentially malicious actors to gain control of a cloud-

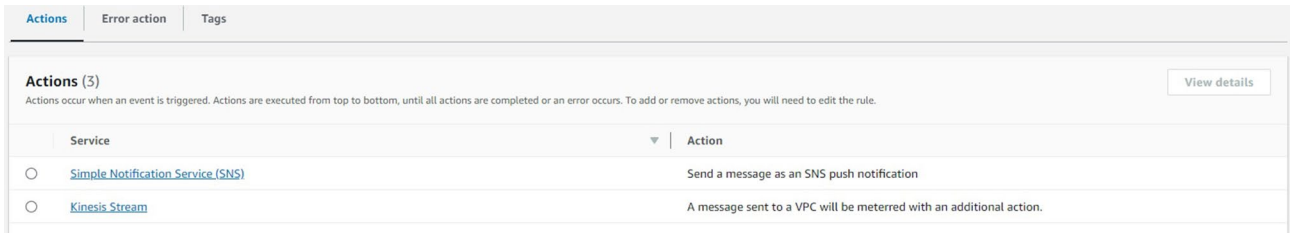


Fig. 5 The messages distributed to AWS SNS and AWS Kinesis

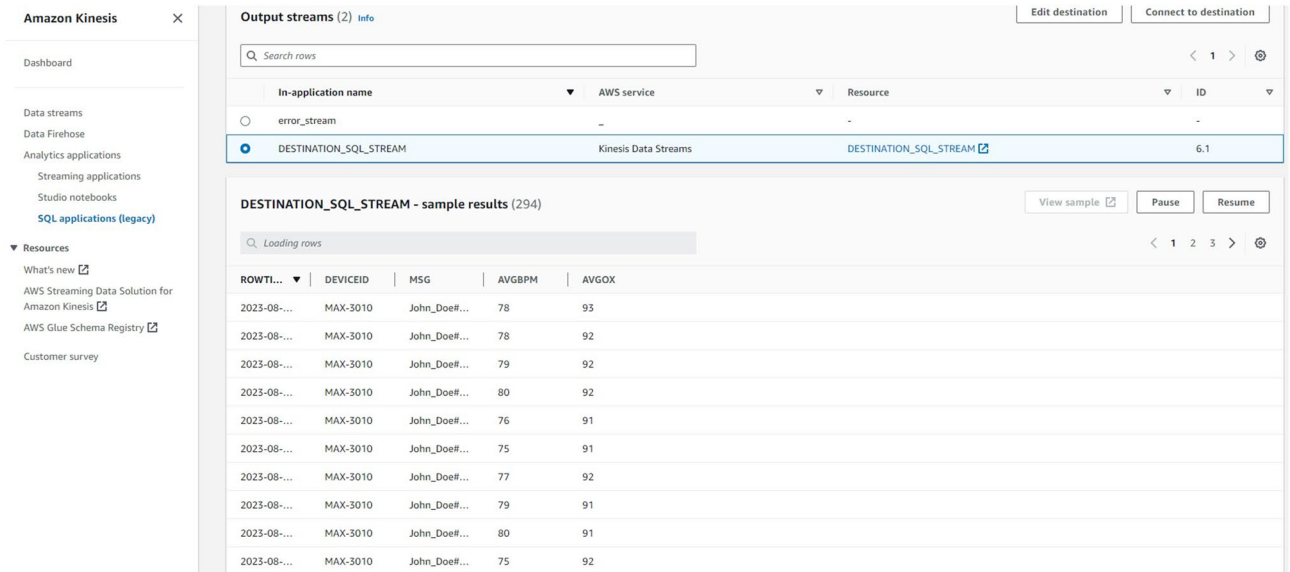


Fig. 6 The output stream after AWS kinesis analytics processing. We can see that the AWS kinesis analytics extracts the data and make the data structured

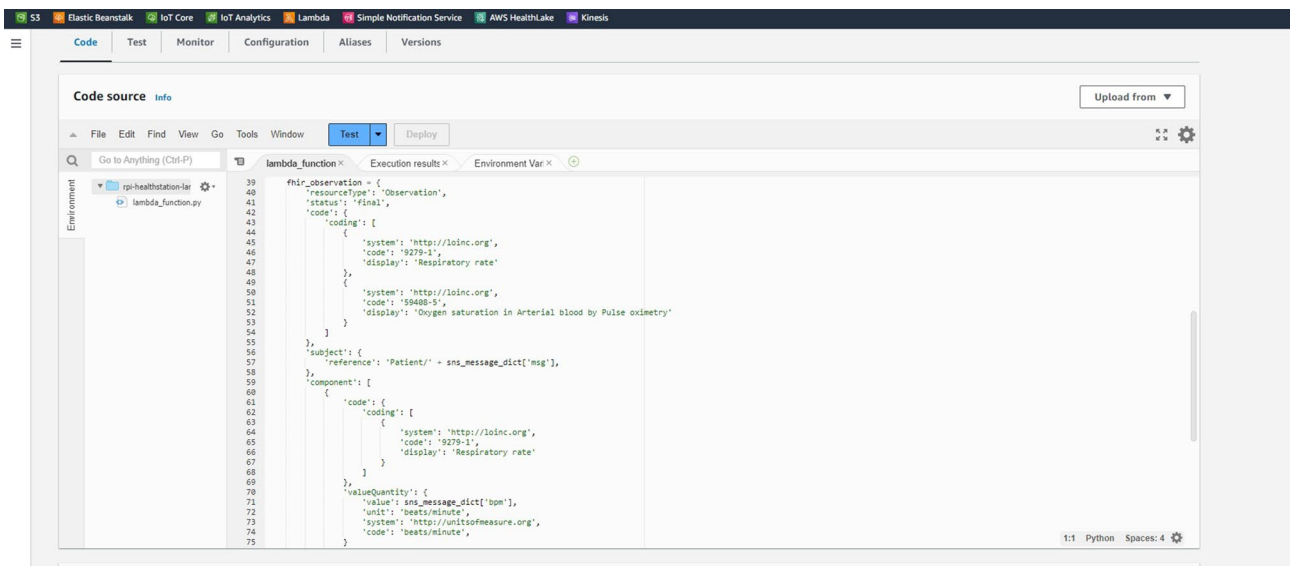


Fig. 7 The Python code of AWS lambda function

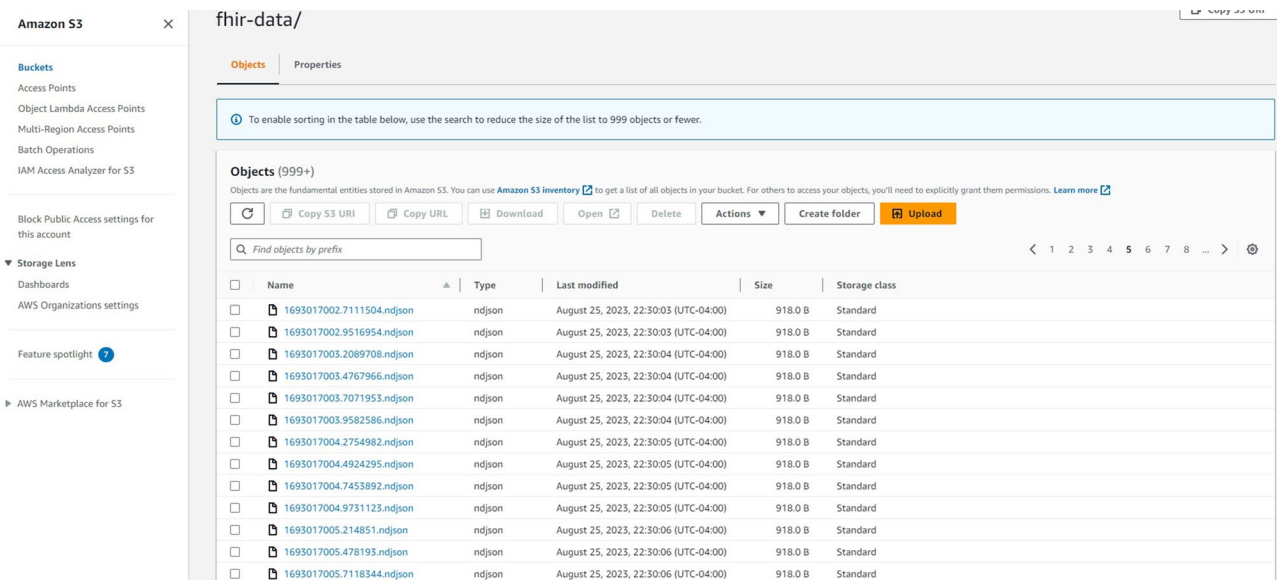


Fig. 8 The real-time data is stored as FHIR files in a AWS S3 bucket

based DT. This issue may be resolved in future work by incorporating Hyperledger Fabric technology.

- The absence of encryption may lead to unintentional personal health data leakage, potentially violating HIPAA regulations. This issue can be addressed by implementing the HTTPS protocol.
- A lack of vendor conformity may impede the adoption of the solution. While DT is a rapidly evolving technology, its maturity has not yet been fully realized. As the technology readiness level of DT becomes more established, this issue is expected to resolve itself.

About the problems of medical data transmission security, various methods have been proposed in literature on how to encrypt data (specifically medical images) before its transmission to cloud. Several algorithms have been tested and implemented to encrypt images and texts to ensure data privacy and security [29–31] before it reaches AWS, that are worth to be considered in future work for complete data security (from a medical system to cloud and vice versa).

About using AWS Healthlake with DT in the environment of emergency healthcare, we discuss as follows. AWS HealthLake is a highly scalable and secure cloud-based data service that is part of the Amazon Web Services (AWS) ecosystem. It is designed to manage, process, and analyze vast amounts of structured and unstructured health data. Its core function is to transform health data from various sources and formats into the standardized Fast Healthcare Interoperability Resources (FHIR) format, which is a globally recognized and utilized standard for health data exchange. Once the data is in the FHIR format, it is much easier to analyze and can

be readily utilized in machine learning and other analytical applications.

The comprehensive nature of AWS HealthLake allows it to store, curate, query, and share data across an entire health-care organization. This includes electronic health records (EHRs), wearable fitness trackers, clinical trials, and other relevant health data sources. HealthLake's data lake architecture and serverless design make it capable of handling petabytes of health data, making it suitable for complex, large-scale health analytics tasks.

In the context of utilizing a Digital Twin (DT) in an emergency healthcare environment, AWS HealthLake can provide a significant edge. The continuous stream of patient data generated by a DT, which replicates the patient's physiological state in real-time, can be processed and stored using AWS HealthLake.

The FHIR-based structured data can be queried to glean valuable insights about the patient's current health status, possible treatment options, and the potential outcomes of different therapeutic interventions. This could include the real-time analysis of vital signs, predicting the body's potential responses to different treatments, or estimating the recovery trajectory given the current interventions being applied.

HealthLake's compatibility with machine learning (ML) models and AI services offered by AWS, such as Amazon SageMaker, allows the development of predictive models that can significantly enhance patient care. Machine learning algorithms could analyze patterns in the patient's DT data to predict the likely progression of their condition. This could range from simple predictions like heart rate or blood pressure

trends, to complex forecasts such as the likelihood of a cardiac event or stroke.

Moreover, these models could be continually refined over time. As the DT continues to generate data, the model could incorporate this new data to improve its accuracy and predictive capabilities. Over time, this could lead to more precise prediction models that can alert healthcare providers to potential health risks before they become critical, contributing to improved patient outcomes.

In a nutshell, AWS HealthLake provides a scalable, secure, and robust data management system for health data. Its use in analyzing and predicting data from a Digital Twin in an emergency healthcare setting can lead to more informed decision-making, timely interventions, and improved patient outcomes. By facilitating the integration of advanced technologies such as machine learning and AI into healthcare, it represents a significant step toward the realization of precision medicine and personalized healthcare.

5 Conclusion

This paper provides an in-depth, critical review of the existing literature at the confluence of digital twins (DT), healthcare, the internet of things (IoT), and cloud computing. These domains, while individually vast and complex, collectively contribute to a rapidly evolving ecosystem that holds significant potential for transformative change in healthcare systems across the globe.

The review encompasses a broad spectrum of technologies that enable the effective functioning of digital twins within the healthcare space. Such technologies range from services provided by cloud platforms that host digital twins, to the intricate data transmission protocols that ensure the seamless transfer of information between digital twins and other systems, and even the storage file formats that help manage and safeguard this crucial data.

Based on this in-depth evaluation and synthesis of the literature, the study proposes a novel, cloud-based digital twin storage solution, specifically designed for emergency healthcare scenarios. Recognizing the pressing need for real-time, accurate, and secure access to patient data during emergencies, the study underscores the development of a practical, viable, and scalable solution. The proposed system leverages state-of-the-art cloud computing technologies, coupled with innovative digital twin applications, to provide healthcare professionals with real-time, actionable insights that can significantly improve patient outcomes.

6 Future Work

Looking toward the future, this paper outlines a roadmap for further research endeavors. One of the main focal points of these endeavors will be refining and optimizing the proposed digital twin storage solution. This refinement process will emphasize stringent adherence to the Health Insurance Portability and Accountability Act (HIPAA) regulations. Given the sensitive nature of healthcare data, ensuring compliance with these regulations is of utmost importance, guaranteeing the security and confidentiality of patient information.

Beyond HIPAA compliance, the future iterations of this solution may incorporate cutting-edge security technologies. One such technology under consideration is the blockchain-based Hyperledger Fabric, a permissioned blockchain infrastructure known for its robust security features and scalability. The incorporation of such advanced security measures would add an additional layer of protection, further enhancing the reliability and trustworthiness of the solution.

The long-term vision for the proposed solution is its evolution into a Digital Twin as a Service (DTaaS) offering. In this scenario, the digital twin technology would be offered as a service over the cloud, enabling healthcare providers to leverage the benefits of digital twins without the need for substantial upfront investment in infrastructure or expertise.

Author Contributions EW designed, programmed, implemented this research, and drafted the manuscript. PT researched on the Sect. 2 and drafted parts of the section. Y-TS conceived this study and instructed the research. The authors read and approved the final manuscript.

Funding No funds, grants, or other support was received.

Availability of Data and Materials Not applicable.

Code Availability Not applicable.

Declarations

Conflict of Interest The authors declare that they have no competing interests.

Ethical Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Grieves M, Vickers J (2017) Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. *Transdisciplinary perspectives on complex systems: new findings and approaches*. Springer, Cham, pp 85–113
- Song Y-T, Qin J (2022) Metaverse and personal healthcare. *Procedia Comput Sci* 210:189–197
- Xie S, Zhu S, Dai J (2021) Feasibility study of intelligent healthcare based on digital twin and data mining. In: 2021 International conference on computer information science and artificial intelligence (CISAI), IEEE, pp 906–911
- Rivera LF, Jiménez M, Angara P, Villegas NM, Tamura G, Müller HA (2019) Towards continuous monitoring in personalized healthcare through digital twins. In: Proceedings of the 29th annual international conference on computer science and software engineering, pp 329–33
- Troyer L, Brady W (2020) Barriers to effective EMS to emergency department information transfer at patient handover: a systematic review. *Am J Emerg Med* 38(7):1494–1503
- Lubin JS, Shah A (2022) An incomplete medical record: transfer of care from emergency medical services to the emergency department. *Cureus*. <https://doi.org/10.7759/cureus.22446>
- Zhang Z, Sarcevic A, Joy K, Ozkaynak M, Adelgais K (2022) User needs and challenges in information sharing between pre-hospital and hospital emergency care providers. In: AMIA annual symposium proceedings, February
- Hu L, Nguyen N, Tao W, Leu MC, Liu XF, Shahriar MR, Sunny NA (2018) Modeling of cloud-based digital twins for smart manufacturing with MT connect. *Procedia Manuf* 26:1193–1203
- Moyaux T, Liu Y, Bouleux G, Cheuet V (2023) An agent-based architecture of the digital twin for an emergency department. *Sustainability* 15:3415
- Sahal R, Alsamhi SH, Brown KN (2022) Personal digital twin: a close look into the present and a step towards the future of personalised healthcare industry. *Sensors* 22:5918
- Kayvanpour E, Mansi T, Hamedani FS (2015) Towards personalized cardiology: multi-scale modeling of the failing heart. *PLoS ONE* 10:e0134869
- Singh S, Shehab E, Higgins N, Fowler K, Reynolds D, Erkoyuncu JA, Gadd P (2020) Data management for developing digital twin ontology model. *Inst Mech Eng* 235:2323–2337
- Kamel Boulos MN, Zhang P (2021) Digital twins: from personalised medicine to precision public health. *J Personal Med* 11:845
- Liu Y et al (2019) A novel cloud-based framework for the elderly healthcare services using digital twin. *IEEE Access* 7:49088–49101. <https://doi.org/10.1109/ACCESS.2019.2909828>
- Martinez-Velazquez R, Gamez R, El Saddik A (2019) Cardio twin: a digital twin of the human heart running on the edge. In: 2019 IEEE international symposium on medical measurements and applications (MeMeA), Istanbul, Turkey, 2019, pp 1–6. <https://doi.org/10.1109/MeMeA.2019.8802162>
- Qi Q, Tao F, Hu T, Anwer N, Liu A, Wei Y, Wang L, Nee AYC (2021) Enabling technologies and tools for digital twin. *J Manuf Syst* 58(Part B):3–21
- Okegbile SD, Cai J, Yi C, Niyato D (2022) Human digital twin for personalized healthcare: vision, architecture and future directions. *IEEE Netw*. <https://doi.org/10.1109/MNET.118.2200071>
- Khaled N, Pattel B, Siddiqui A (2020) Digital twin development and deployment on the cloud: developing cloud-friendly dynamic models using Simulink@SimscapeTM and Amazon AWS. Academic Press, Cambridge
- Saad A, Faddel S, Mohammed O (2020) IoT-based digital twin for energy cyber-physical systems: design and implementation. *Energies* 13(18):4762
- Wang Z, Gupta R, Han K, Wang H, Ganlath A, Ammar N, Tiwari P (2022) Mobility digital twin: concept, architecture, case study, and future challenges. *IEEE Internet Things J* 9(18):17452–17467
- Saraswat M, Tripathi RC (2020) Cloud computing: comparison and analysis of cloud service providers-AWs, Microsoft and Google. In: 2020 9th international conference system modeling and advancement in research trends (SMART), Moradabad, India, 2020, pp 281–285. <https://doi.org/10.1109/SMART50582.2020.9337100>
- Amazon Web Services Inc. (2022) Architecting for HIPAA security and compliance on Amazon web services. [https://docs.aws.amazon.com/whitepapers/latest/architecting-hipaa-security-and-compliance-on-aws.html](https://docs.aws.amazon.com/whitepapers/latest/architecting-hipaa-security-and-compliance-on-aws/architecting-hipaa-security-and-compliance-on-aws.html). Accessed 10 June 2023
- Microsoft Corporation (2022) Introduction to Azure security. <https://learn.microsoft.com/en-us/azure/security/fundamentals/overview>. Accessed 10 June 2023
- Google LLC (2022) Google Cloud security foundations guide. <https://services.google.com/fh/files/misc/google-cloud-security-foundations-guide.pdf>. Accessed 10 June 2023
- Đorđević BS, Jovanović SP, Timčenko VV (2014) Cloud computing in Amazon and Microsoft Azure platforms: performance and service comparison. In 22nd telecommunications forum TELFOR (TELFOR), Belgrade, Serbia, 2014, pp. 931–934, <https://doi.org/10.1109/TELFOR.2014.7034558>
- Google LLC (2023) Compute Engine pricing. <https://cloud.google.com/compute/all-pricing>. Accessed 10 June 2023
- Amazon Web Services (2023) AWS pricing. https://aws.amazon.com/pricing/?aws-products-pricing.sort-by=item.additional.Fields.productNameLowercase&aws-products-pricing.sort-order=asc&awsf.Free%20Tier%20Type=*all&awsf.tech-category=*all. Accessed 10 June 2023
- Muhammed AS, Ucuz D (2020) Comparison of the IoT platform vendors, Microsoft Azure, Amazon Web Services, and Google Cloud, from users' perspectives. In: 2020 8th International symposium on digital forensics and security (ISDFS), Beirut, Lebanon, 2020, pp 1–4. <https://doi.org/10.1109/ISDFS49300.2020.9116254>
- Denis R, Madhubala P (2021) Hybrid data encryption model integrating multi-objective adaptive genetic algorithm for secure medical data communication over cloud-based healthcare systems. *Multimedia Tools Appl* 80:21165–21202
- Razzaq MA, Shaikh RA, Baig MA, Memon AA (2017) Digital image security: fusion of encryption steganography and watermarking. *Int J Adv Comput Sci Appl* 8(5):224–228
- Bairagi A, Khondoker R, Islam R (2016) An efficient steganographic approach for protecting communication in the internet of things (IoT) critical infrastructures. *Inf Secur J Glob Perspect* 25:1–16. <https://doi.org/10.1080/19393555.2016.1206640>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.