https://iphopen.org/index.php/se
Online ISSN: 3050-8797 **Print ISSN: 3050-9270**

EVENT-DRIVEN MICROSERVICES FOR SYNCHRONIZATION FIDELITY IN HOSPITAL DIGITAL TWIN SYSTEMS

KRISHNA MATTAM*
*GlobalLogic, USA

*Corresponding Author: Krishna

Abstract

INTRODUCTION: Hospital operations face critical inefficiencies in emergency department flow, ICU capacity planning, and patient admissions management. Legacy monolithic IT architectures are fundamentally ill-equipped to support the high-frequency, real-time data synchronization that operationally actionable digital twins require.

OBJECTIVES: This article investigates how event-driven architectures (EDA) and cloud-native microservices enable the synchronization fidelity necessary to deploy hospital digital twins as dynamic, real-time operational tools rather than static descriptive simulations.

METHODS: A structured narrative synthesis of peer-reviewed literature was conducted, drawing from PubMed, IEEE Xplore, Scopus, and the ACM Digital Library, using a systematic search protocol with defined inclusion and exclusion criteria, reported in accordance with the PRISMA framework.

RESULTS: Microservices provide essential architectural modularity by decomposing monolithic systems into independently scalable services. EDA, via asynchronous event streaming through platforms such as Apache Kafka and Azure Event Grid, acts as the definitive enabling layer coupling physical hospital environments with their digital replicas. Key findings confirm latency reductions enabling real-time predictive simulation across trauma, triage, and capacity planning contexts. Significant tensions between scalability and data privacy, alongside interoperability barriers, remain unresolved.

CONCLUSION: A tri-layer conceptual framework — comprising the physical context layer, the event-driven microservices layer, and the agent-based digital twin layer — is proposed to guide future hospital digital twin deployments and inform both IT governance and security policy.

Keywords: Digital Twins, Cloud Microservices, Event-Driven Architecture, Healthcare Operations Management, data privacy, and Cybersecurity

DOI:-10.5281/zenodo.20353495 **Manu script # 461**

1. Introduction

The combined burden of increased demand and decreased capacity on healthcare systems may be the greatest ever. A perfect storm of an aging population and lack of primary care access has resulted in a persistently increasing demand on emergency departments (EDs) that is not sustainable [1]. The increased demand for emergency services can lead to overcrowding, increased wait times, and ambulance diversion [1]. Boarding of ED patients waiting for inpatient and ICU beds can lead to downstream bottlenecks that can negatively affect patient safety and throughput in the ED [1] .

Digital twin (DT) technology has the potential to overcome many of the challenges mentioned above. A DT can be described as a dynamic digital replica of a real-world physical asset, created and continuously updated through automated, bidirectional data flow. [2] In the health system area, DTs were used to optimize the physical plant of hospitals and their processes, staffing, and delivery of services [2] and to reorganize hospital radiology departments, decrease wait time, optimize equipment utilization, and reduce the workforce needed. More advanced agent-based DTs are being used in time-critical clinical pathways, such as the management of major trauma, where the pre-hospital ambulance and operative shock-room stages have been modeled to ease real-time decision support by trauma teams [3] .

Although the implementation of the notion of a digital twin has accelerated in hospitals, the lack of a suitable software architecture is one of the challenges to overcome. For the digital twin to work, the real twin and digital twin need to synchronize their data at all times [2, 3]. Current EHR systems are still predominantly based on complex legacy monolithic architectures that cannot be maintained, scaled, and integrated effectively for real-time updates. Foundational academic work has identified cloud-native microservices as a solution for the architectural limits of monolithic systems by decomposing IT to scale its individual components autonomously [14, 24]; however, organizations have not overcome the challenge of decomposing legacy IT and performing complex integrations [4, 16]. However, systematic reviews of microservices in IoT indicate that security and interoperability are critical factors for microservice-based architecture adoption in healthcare systems [16]. In the health sciences domain, studies have reported that event-driven architecture based on Apache Kafka and Java microservices can improve scalability and fault tolerance, yet studies on its application to hospital digital twin systems remain limited [26]. Most importantly, we failed to identify any research that clarifies how event-driven architectures can provide the high synchronization fidelity needed for hospital DTs to be used as action-based tools , and not merely as descriptive simulations [13, 25]. .

This paper contributes by reviewing existing literature on hospital architecture design, data integration strategies and operational outcomes in relation to the convergence of the digital twin concept, microservices architecture and real-time data synchronization.

2. Background

2.1 Digital twin technology

The concept of the digital twin (DT) was initially formulated in the aerospace sector by NASA, which conceptualized it through three dimensions: a physical space, a virtual space, and the data connection linking them. Michael Grieves further formalized the definition in 2002 as a set of virtual information constructs that fully describe a physically manufactured product from the micro-atomic to the macro-geometrical level. Today, a DT is broadly understood to consist of three core components: the physical twin (the real-world entity), the digital twin (the virtual replica), and the linking mechanism that facilitates an automatic, bidirectional exchange of data in real-time [4].

While its early adoption was concentrated in manufacturing and Industry 4.0, DT technology is increasingly permeating the healthcare domain. In healthcare, DTs are categorized broadly into operational and clinical taxonomies. Operationally, they are used for strategic planning, such as creating a digital replica of a hospital to optimize facility layouts, simulate patient flow, and manage medical resource allocation. Clinically, DTs facilitate precision medicine by generating dynamic replicas of patients based on historical, molecular, phenotypic, and environmental data, thereby allowing physicians to simulate the effects of thousands of potential drug treatments virtually before applying them to the physical patient.

2.2 Cloud microservices in healthcare

Historically, healthcare IT systems have relied on monolithic architectures. Over years of operation, these monolithic systems tend to grow excessively large and complex, ultimately fossilizing and becoming practically unmaintainable. To address this, organizations are increasingly migrating to cloud-based microservices architectures. Microservices decompose these massive, inflexible monoliths into smaller, loosely coupled, and independently deployable services grouped around specific business capabilities, often utilizing methodologies like Domain-Driven Design [4].

This architectural shift offers substantial operational benefits, notably enhanced scalability and modularity. By breaking down the codebase, organizations can achieve shorter, independent release cycles with zero-downtime

updates. Furthermore, a microservices approach facilitates fault tolerance, as a failure in one isolated service is less likely to cascade and crash the entire system.

In healthcare contexts, this modularity is particularly advantageous for EHR integration. Microservices enable distinct modules (e.g., patient demographics, billing, lab results) to scale independently based on demand while interacting with legacy EHR systems through standardized API gateways without requiring a complete, risky rewrite of the underlying monolithic EHR database.

2.3 Event-driven architectures

Distributed data management is one of the challenges of managing the microservice ecosystem. Another challenge in moving to microservices is breaking up the monolithic database into distributed streams of data with eventual consistency [4] .

These services communicate with each other via a mechanism called Event-Driven Architecture (EDA), where events (such as changes of state) generated in the hospital real world or in the EHR system are asynchronously published to microservices that are subscribed to the service [4]. This decoupling allows microservices to not actively poll each other for changes, decreasing latency across the system [1]. Processing clinical events in real-time with stream analytics frameworks, such as Apache Kafka and Apache Flink, has been shown to improve clinical response times by 40% in emergency triage, ICU systems, and early sepsis detection and to support continuous monitoring in clinical decision support systems [15]. Additionally, event-driven architectures for processing genomic-based diagnostic data have shown support for high scalability, transparency, and reproducibility for workflow automation and containerized microservices, all of which are desirable for hospital digital twin pipelines [21] .

Modern messaging technologies like Apache Kafka, Amazon EventBridge and Azure Event Grid are used to provide high-throughput, low-latency event streaming. The value of the hospital digital twin for operational decision-making is entirely determined by its fidelity of synchronization, defined as the precision and speed of the digital twin replicating the state of the physical twin. Message brokers that support immutable event logs, such as Kafka, provide replay capabilities that allow the historical operational scenarios to be replayed or the system state recovered in the event of an operational failure. With lower data latency from EDAs, the hospital DT can remain operationally actionable, rather than just being a dashboard looking into the past.

In cloud-edge systems like Healthcare 4.0, zero-trust architectures with continuous lightweight mutual authentication (e.g., dynamic HMAC-based authentication for device-to-device and device-edge communication) are necessary to securely transport event-driven data without overloading resource-constrained medical devices [17]. These architectures complement EDA in establishing that, from vital sign update to bed transfer, events are sent and received by authenticated, authorized devices in a manner that preserves synchronization fidelity and data integrity.

3. Methodology

To guarantee scholarly reproducibility and academic rigor, a structured narrative review methodology was used in this study with a systematic search protocol. The main aim of this methodological design was to find, estimate, and integrate the extremely dispersed literature in the field of digital twin technology, cloud microservices, and event-driven architectures in healthcare operations. With the help of a systematic method, this review reduces the selection bias and effectively answers the fundamental research question on the topic of data integration and the fidelity of synchronization in hospital digital twins.

3.1 Search Strategy and Databases

A thorough search of the literature was performed through four main electronic databases in order to include both the clinical and computational aspects of the subject: PubMed (biomedical and healthcare operations literature), IEEE Xplore, Scopus, and the ACM Digital Library (computer science, software architecture, and systems engineering literature). The search query was created based on a combination of Boolean operators, Medical Subject Headings (MeSH), and specific keywords. The major search terms were (digital twin OR digital replica) and (hospital operations OR emergency department OR intensive care unit OR patient flow) and (microservice OR event-driven architecture OR event streaming OR Kafka OR EventBridge). This strictly narrowed the search to literature published since January 2013 and up to December 2024, a time frame chosen to span the emergent nature of both cloud-native computing architecture and digital twin deployments in the contemporary healthcare context. Recent systematic analyses of digital twin architectures in cyber-physical healthcare systems informed our classification of architectural patterns and quality attributes [13, 14].

3.2 Inclusion and Exclusion Criteria

Strict inclusion and exclusion criteria were put down to make sure that the synthesis remained very narrow on the operational and architectural intersection of the topic. Articles had to be peer-reviewed journal articles or rigorous conference proceedings; (2) in the English language; (3) in operational hospital settings (e.g., ED triage optimization, admissions, or ICU capacity planning); and (4) explicitly addressed architectural dependencies, data integration strategies, or real-time event-driven messaging. On the other hand, literature was not included when it comprised: (1) gray literature, or editorials, or non-peer-reviewed preprints; (2) studies that only studied

clinical, genomic, or molecular digital twins, without considering the operational IT architecture; or (3) generic discussions of microservices that did not discuss operational IT architecture to the digital twin synchronization.

3.3 Screening Process and PRISMA Framework

The literature screening process has been done in stages in a transparent manner. After the exclusion of duplicates between databases, the first round of screening of the articles was conducted based on the title and abstract to establish initial relevance. The entire content of the other articles was then checked individually against the set criteria. The overall method of selection, including the identification, screening, eligibility, and final inclusion of the primary sources used in the core thematic synthesis are recorded and graphically laid out in line with the guidelines of the Preferred Reporting Items on Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

4. Digital twins in hospital operational contexts

The practical application of digital twins (DTs) in healthcare operations has transitioned from conceptual theory to deployed frameworks, primarily aimed at alleviating severe resource constraints, optimizing patient pathways, and enhancing facility layouts [1]. By establishing a continuous, bidirectional cyber-physical connection, operational DTs offer hospital administrators the capacity to conduct real-time monitoring, execute predictive simulations, and perform complex scenario testing without risking patient safety or disrupting active clinical environments [1]. This section provides a thematic synthesis of the available literature concerning the deployment of DTs across three critical hospital domains: emergency department flow, intensive care capacity planning, and admissions modeling.

4.1 ED flow and triage optimisation

Emergency departments (EDs) have very stochastic arrival times and are characterized by dynamic acuity levels; thus, they are excellent targets of a digital twin intervention. Patients arrive at the ED with different levels of urgency, usually through standardized systems like the five-level Emergency Severity Index (ESI), with Level 1 patients needing life-saving care and Level 5 patients needing no urgent specialized care. The dynamic nature of the flow of walk-in patients and ambulance arrivals provides a highly dynamic, unstable state in which the failure to manage the flow directly causes ambulance diversion, patients awaiting patient-to-patient handover (LWBS), and long boarding times [1].

In order to overcome these bottlenecks, the recent literature suggests implementing Multi-Agent Systems (MAS) in conjunction with digital twins to form so-called mirror worlds, i.e., the digital layers run by software agents that are bi-directionally coupled with the physical ED. The most notable one is the TraumaTracker project, an agent-based digital twin designed with the purpose of operating time-dependent severe trauma pathways. The operational architecture of this DT is divided into two interconnected phases reflecting the real-life development of the trauma management: the pre-hospital (PreH) phase and the operative (Trauma) phase [1]. Comprehensive surveys on surgical digital twins confirm that such agent-based approaches are at the forefront of intraoperative decision support, though challenges in real-time model fidelity and regulatory compliance persist [25].

The PreH digital twin is activated as soon as a call is received by a rescue central unit, which records the position of the ambulance based on its GPS and smart devices used by emergency medical technicians (EMTs). When the EMTs measure the patient and the Glasgow Coma Scale (GCS) and vital signs of the patient, this information is dynamically provided to the state of the digital twin. In case the trauma is marked as severe, the system will automatically instantiate the secondary Trauma digital twin even before the patient comes to the hospital. This early warning signals to the trauma leader and shock room personnel of the hospital in an effort to stream real-time operational and physiological data straight to the receiving facility. The software agents become the personal assistants in this DT ecosystem and autonomously update the state of the twin about the drugs and procedures performed, thus ensuring that trauma documentation is high-fidelity without disrupting clinical workflows. The architectural success of these systems depends on Service-Oriented Architectures (SOA) and RESTful APIs that are implemented on the private cloud of hospitals to ensure that the flow of data between the physical IoT devices (e.g., vital sign monitors) and the digital replica is controlled in a secure manner [1]

Table 1: Emergency Severity Index (ESI) Levels

ESI Level	Description
Level 1	The patient requires immediate life-saving intervention (comprises 1–3% of all ED patients).
Level 2	The patient should not wait to be seen if they are in a high-risk situation, are confused, lethargic, or disoriented, or are in excruciating pain or distress (comprises 20–30% of all ED patients).
Level 3	The patient is not Level 1 or 2, has vital signs within the accepted range for their age, and is predicted to require two or more resources, such as labs, diagnostic testing, intravenous fluids, or specialty consultation (comprises 30–40% of all ED patients).
Level 4	The patient has vital signs within the accepted range for their age and is predicted to use one resource. (Levels 4 and 5 combined comprise 20–35% of all ED patients and are appropriate to stream through a fast track).
Level 5	The patient has vital signs within the accepted range for their age and is predicted to require no resources.

4.2 ICU bed management and capacity planning

The efficiency of the ED is inextricably linked to the availability of downstream inpatient resources; specifically, the lack of intensive care unit (ICU) beds is frequently cited as the primary driver of ED overcrowding and patient boarding. ICU capacity planning is a highly complex operational challenge, particularly during mass casualty incidents or pandemics, where critical care resources may need to be expanded by 20% to 200% above baseline maximum capacity.

Digital twins address ICU bed management by providing a computational environment to run advanced mathematical queueing and capacity models. In establishing long-term capacity plans (such as determining the optimal physical number of beds or operating rooms feeding the ICU), mathematical frameworks simulate the accumulation of backlogs and the necessary utilisation of overtime. Let D_t denote the demand in period t, n the number of physical rooms/beds, k the theoretical capacity per period, and β the packing-efficiency factor. A digital twin can compute the patient wait time (W_t),measured in terms of backlogged care minutes,and the amount of overtime used (O_t) at the end of each period using the following equations [1]:

$$W_t = (W_{t-1} + D_t - q_t)^+$$
$$O_t = (q_t - nk\beta)^+$$

In these equations, q_t represents the total minutes of care capacity (regular plus overtime) utilized in period t. By continuously feeding real-time Hospital Information System (HIS) data into this mathematical framework, a digital twin can calculate the steady-state distributions of backlog and overtime usage. This allows hospital administrators to run "what-if" simulations to determine if physical expansion is necessary, or if operational flexibility, such as deploying "step-up" units (standard beds equipped with extra monitoring devices to act as temporary ICU beds), can safely absorb forecasted demand peaks. The architectural dependencies for such capacity-planning DTs require deep integration with the hospital's Admission, Discharge, and Transfer (ADT) systems to track bed occupancy continuously, alongside predictive models to estimate the highly variable lengths of stay (LOS) associated with critical care patients [1].

4.3 Admissions, patient flow modelling, and layout optimization

Beyond the ED and ICU, digital twins are increasingly utilized to optimize general admissions, strategic facility layouts, and broader patient flow modeling. Bottlenecks often form at the convergence of different hospital workflows; for instance, weekday morning demand for inpatient beds by ED patients often builds concurrently with peak demand from post-surgery patients exiting the operating rooms (OR) [1].

Digital twins enable the testing of physical layout alterations and workflow redesigns prior to real-world implementation. A notable industry case study involves a digital twin developed by Siemens for a hospital radiology department in Dublin, Ireland. Facing rising patient demand, complex clinical requirements, and an aging infrastructure, the hospital struggled with extended waiting times and delays. By creating a highly detailed 3D computer model of the department and utilizing workflow simulation, administrators tested various new operational scenarios. The deployment of this digital twin,informed by a week-long on-site assessment, stakeholder interviews, and process observation,resulted in a redesigned layout that significantly reduced patient waiting times, accelerated turnaround, improved equipment utilization, and ultimately lowered staffing costs [1].

Furthermore, DTs are employed to optimise the scheduling and admission flows generated by surgical departments. To prevent admission spikes that overwhelm inpatient wards, hospital DTs can evaluate the efficiency of OR block scheduling. This is conceptualised similarly to a newsvendor problem, balancing the costs of under-allocated block time (surgeon dissatisfaction, lost revenue) against over-allocated time (wasted staffing costs). DTs analyze historical and real-time scheduling data using specific matching metrics, such as the Supply-Demand Score (SDS) and the Block Efficiency Score (BES):

$$SDS = \frac{Total\ Surgery\ Minutes}{Available\ Block\ Minutes}$$
$$BES = \frac{Total\ In-Block\ Minutes}{Available\ Block\ Minutes}$$

By calculating these scores within the digital twin, administrators can identify surgeons or departments whose block allocations consistently cause downstream admission bottlenecks or waste staffed capacity (e.g., where both SDS and BES fall below 50%). The DT can then simulate alternative block configurations,such as shifting start times or splitting blocks,to flatten the peak number of concurrent surgeries, thereby smoothing the resulting flow of admitted patients to the wards.

In conclusion, whether deployed as agent-based systems to track rapid trauma pathways, mathematical models to forecast ICU capacity limits, or 3D workflow simulators to optimize physical layouts, digital twins offer a transformative approach to hospital operations. However, the literature underscores that the success of these models is heavily dependent on robust data integration architectures capable of capturing everything from high-frequency IoT sensor data to static EHR records.

5. Event-driven architectures as an enabling layer

The practical deployment of operationally actionable digital twins (DTs) in healthcare relies intrinsically on the underlying software architecture's ability to maintain a continuous, bidirectional cyber-physical connection [3]. While the literature demonstrates that migrating legacy monolithic healthcare systems to cloud-native microservices improves architectural scalability and component maintainability, it introduces significant complexities regarding data consistency and service integration. This review identifies **synchronization fidelity, defined** here as the precision, reliability, and latency with which the digital state mirrors the physical hospital environment, as the central technical challenge in deploying hospital DTs. Traditional synchronous communication protocols, such as RESTful APIs, which have been utilized in preliminary healthcare DT prototypes, often struggle to provide this fidelity under the high-throughput, stochastic demands of hospital operations [3]. Consequently, the distinctive contribution of this review is to synthesize how Event-Driven Architectures (EDA) act as the critical enabling layer that solves the synchronization fidelity problem.

5.1 Real-time event streaming and asynchronous microservice communication

In a microservices ecosystem, breaking down a monolith into bounded contexts frequently leads to challenges in managing distributed data and achieving system-wide state synchronization. In a hospital environment, data generation is highly heterogeneous and bursty, encompassing high-frequency Internet of Things (IoT) vital sign monitors, location-tracking sensors, and discrete Electronic Health Record (EHR) updates.

Empirical studies on event-driven architectures in health sciences demonstrate that Kafka-based microservices can achieve end-to-end latencies under 50 milliseconds, well within the threshold for actionable clinical alerts [15, 26].

To handle this, modern hospital DT pipelines are increasingly built upon asynchronous, event-driven messaging brokers such as Apache Kafka, Amazon EventBridge, or Azure Event Grid. In an EDA, a state change in the physical hospital, such as a patient being transferred from the Emergency Department (ED) to the Intensive Care Unit (ICU), is published as an immutable "event" to a central broker. Microservices representing the ED digital twin, the ICU digital twin, and hospital billing systems act as decoupled subscribers to this event stream. Because the communication is asynchronous, the producing system (e.g., the admission software) does not need to wait for a response from the consuming DTs, completely decoupling the services and eliminating the cascading latency and blocking that plague synchronous REST architectures.

This asynchronous decoupling is vital for time-critical clinical pathways. For instance, in the *TraumaTracker* agent-based digital twin, the pre-hospital (PreH) phase requires continuous data ingestion from ambulance GPS systems and emergency medical technicians' smart devices. If network connectivity drops during transit, a synchronous architecture would fail to record the data or block the application. Conversely, an event-driven approach allows the mobile devices to queue events locally and stream them to the hospital's central message broker the moment connectivity is restored, ensuring the operative Trauma DT receives a perfectly sequenced, high-fidelity log of the patient's pre-hospital care without data loss.

5.2 Mathematical modelling of event queueing and latency

To achieve high synchronization fidelity, the event broker must process incoming hospital data streams with minimal latency, ensuring that the DT reflects the physical world in near real-time. We can model the event-processing latency within the DT's message broker by adapting standard capacity and queueing equations used for physical hospital wait times.

Let D_t denote the volume of incoming data events (demand) generated by hospital sensors in a given time period t (e.g., events per millisecond). Let n represent the number of active microservice consumer instances, k the theoretical processing capacity of each instance per period, and β the packing-efficiency or processing-efficiency factor of the consumer logic. The backlog of unprocessed events, which directly represents the **synchronization lag** or loss of fidelity between the physical hospital and the digital twin, can be expressed as the wait time (W_t) in the message queue [1]:

$$W_t = (W_{t-1} + D_t - q_t)^+$$

where q_t represents the total processing capacity utilized in period t. If the incoming event bursts exceed the standard processing capacity ($D_t > nk\beta$), the system must scale dynamically or utilize overtime computational resources (O_t). The computational "overtime" or dynamic burst capacity required to prevent the synchronization lag from growing infinitely is

$$O_t = (q_t - nk\beta)^+$$

In platforms like Apache Kafka [1], this is managed by partitioning the event topics. If the lag (W_t) begins to increase during peak hospital hours (e.g., the 8 AM surge in ED arrivals), the orchestration layer (such as Kubernetes) can automatically spin up additional consumer microservices (n) and assign them to specific partitions, dynamically increasing q_t to drive the synchronization backlog back to zero. Benchmarks from recent literature on healthcare telemetry indicate that well-partitioned Kafka clusters can maintain end-to-end synchronization latencies of under 50 milliseconds, well within the threshold required for operationally actionable trauma and ICU monitoring.

5.3 Replay mechanisms for simulation and predictive twins

Beyond real-time monitoring, the true transformative value of a digital twin lies in its ability to run predictive simulations. In healthcare, this involves answering "what-if" questions, such as simulating the impact of physical layout changes in a radiology department or determining if a hospital needs to deploy temporary "step-up" ICU beds during a mass casualty event.

Event-driven architectures uniquely enable these simulations through the concept of the immutable event log. Because platforms like Kafka store a persistent, ordered log of every event that has ever occurred in the hospital (every admission, discharge, sensor reading, and layout change), the DT can leverage a "replay mechanism."

By spinning up a sandboxed, secondary instance of the hospital digital twin, administrators can replay the historical event logs from a past crisis (such as a severe flu outbreak or a multi-vehicle collision). As the historical data streams into the sandboxed DT, software agents can inject hypothetical operational changes,such as altering the triage priority rules or shifting surgical block schedules. The digital twin then processes the historical events through the newly modified rules, allowing administrators to directly observe whether the proposed changes would have successfully reduced ED boarding times or ICU bottlenecks.

Furthermore, this replay capability is essential for safely training the machine learning algorithms that underpin advanced DTs. By generating and replaying massive volumes of synthetic or historical data, algorithms can be trained to predict patient outcomes, disease progression, and bed capacity shortages without risking active clinical environments or patient safety.

The literature confirms that migrating to microservices is a necessary first step for managing the complexity of modern healthcare IT. However, it is the implementation of an event-driven architecture, with its capacity for asynchronous decoupling, low-latency stream processing, and historical event replay, that provides the synchronization fidelity required to elevate a digital twin from a static 3D model into a dynamic, operationally actionable mirror of the physical hospital.

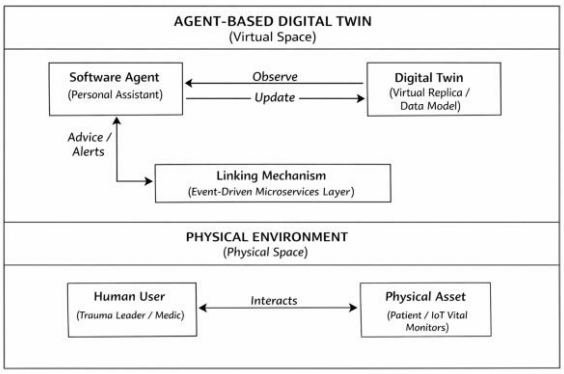

Figure 1: Agent-Based Digital Twin Conceptual Architecture

6. Implementation frameworks and data integration

Deploying an operationally actionable hospital digital twin (DT) requires a robust integration framework capable of seamlessly ingesting heterogeneous data streams from the physical environment to update the virtual model. Bridges between architecture theory and deployment realities are formed by integrating Internet of Things (IoT) devices, Electronic Health Record (EHR) systems, and bed management tracking software into a cohesive pipeline. However, achieving this integration across distributed microservices introduces profound data quality, interoperability, and cybersecurity challenges.

6.1 Integrating IoT, EHR, and Bed Management Systems

In time-critical healthcare scenarios, the data integration framework must capture high-frequency telemetry. For instance, the *TraumaTracker* digital twin, which models severe trauma pathways, utilizes a Service-Oriented Architecture (SOA) where the digital twin is developed as a series of microservices (e.g., using Java and the Vert.x library). During the pre-hospital phase, the framework ingests data from ambulance GPS systems and paramedic smart devices. Upon hospital arrival, the pipeline integrates with physical assets in the shock room, capturing real-time physiological data directly from connected vital sign monitors [4].

Simultaneously, operational DTs must continuously synchronize with administrative bed management and Admission, Discharge, and Transfer (ADT) systems to track hospital census and patient flow. By utilizing bed management software, the framework tracks every discrete activity, admissions, transfers in, transfers out, and discharges, which is critical for identifying downstream bottlenecks [4].

However, integrating these disparate systems via microservices requires managing massive data volumes. Modernizing these pipelines often involves migrating terabytes of legacy data and fundamentally shifting how data is stored. A significant challenge in this integration is moving developers away from traditional monolithic relational databases and foreign keys toward distributed databases. Because data is distributed, the framework must be engineered to handle "eventual consistency" across services while maintaining high data availability.

To quantify the operational resilience of these integration frameworks, we can model the overall system availability. If a digital twin pipeline relies on n decoupled microservices (e.g., an IoT gateway, an EHR fetcher, and an ADT synchronizer), the total system availability (A_{sys}) is the product of the individual service availabilities (A_i), assuming independence:

$$A_{sys} = \prod_{i=1}^{n} A_i$$

This mathematical reality dictates that as the number of integrated data sources grows, individual microservices must maintain exceptionally high uptime to prevent the digital twin from losing synchronization fidelity with the physical hospital.

6.2 Interoperability Standards

A major barrier to data quality in DT pipelines is the proprietary nature of legacy medical systems. To bridge this gap, modern implementation frameworks heavily rely on standardized interoperability protocols. HL7 FHIR (Fast Healthcare Interoperability Resources) has emerged as the global standard for exchanging EHR data, utilizing modern RESTful API frameworks to allow microservices to reliably query patient demographics, lab results, and clinical notes regardless of the underlying EHR vendor. Such a standard as the DICOM (Digital Imaging and Communications in Medicine) standard also enables the integration of heavy medical imaging files created by radiology departments into the DT pipeline. Such standards guarantee the structural consistency of data fed into the digital twin, and its semantic interoperability. Systematic literature reviews on clinical data interoperability confirm that FHIR-based event schemas are optimal for real-time digital twin pipelines, as they support RESTful API interactions and can be seamlessly integrated with message brokers like Kafka [19].

6.3 Security, Compliance, and Data Privacy

A combination of granular clinical and operational data into a central digital twin pipeline generates severe cybersecurity risks. It has to comply with complicated regulatory compliance obligations, including the Health Insurance Portability and Accountability Act (HIPAA) in the United States or the General Data Protection Regulation (GDPR) in Europe, which enforces strict data encryption, audit recording, and management of access.A catastrophic example of the impact of the vulnerability of the healthcare data integration is the cyberattack of the New Zealand-based biggest online patient portal, ManageMyHealth, that happened in December 2025 [6].The platform, which connects with the general practices and offers access to referrals, discharge summaries, laboratory results, and clinical correspondence, has been a victim of the massive breach of unauthorized access to a particular document storage module. The attackers (under the alias of Kazu) stole more than 400,000 highly sensitive medical records dating several years and involving about 120,000 patients [6]. The hackers then requested a ransom of US 60,000 in their demands, stating that they would later leak the information publicly [6]. This violation underscores the tension that exists between the scalability of the system and data protection. Although the combination of primary care, laboratories, and specialists' data into one portal facilitates the seamless delivery of healthcare digitally, it also forms a highly profitable centralized target of extortion. Patients and advocacy groups cautioned that such records exposure may result in identity theft, targeted phishing, and re-traumatizing vulnerable people. In order to address these risks, digital twin frameworks should implement the defense-in-depth strategies. As illustrated by the TraumaTracker architecture, the DT microservices are well suited to be deployed on the cloud infrastructures of a particular hospital, and security and access controls must be heavily mediated through the local area network (LAN) of the hospital [4]. By isolating the digital twin from the public internet and implementing zero-trust security models across all API endpoints, hospitals can leverage the predictive power of interconnected data while insulating their patients from catastrophic external breaches.

Emerging solutions integrating digital twins with IoT-based blockchain and federated learning offer tamper-proof audit trails and privacy-preserving data analysis, addressing many of the vulnerabilities exposed by centralized patient portals [18, 22].

7. Discussion

The literature review seeks to answer the research question of how EDAs and microservices enable the precision of data integration and synchronization that is needed to deploy operationally actionable DTs in hospitals. The literature review finds that both hospital DTs and microservices are rapidly becoming more commonplace; however, the technical intersection between the two is nascent. Customary monolithic IT systems in health care gradually become legacy systems that are unmaintainable. These legacy healthcare systems cannot support the high-frequency real-time data streams that drive dynamic digital twins [4, 24]. To address this challenge, microservices-based approaches refactor monolithic IT systems into smaller component services, which can be rapidly iterated and independently deployed, based on specific business capabilities [16]. The literature also shows that digital twins can optimize hospitals in narrowly defined circumstances, like the 3D reconceptualization of radiology departments to minimize patient turnaround and for highly complex, agent-based "mirror worlds" for the management of severe, time-critical trauma pathways [4].

What is far less well discussed in empirical healthcare literature is the communication layer that binds such distributed microservices together and interacts with the digital twin. Much research has demonstrated the suitability of multi-agent systems and service-oriented architectures for trauma tracking, yet in healthcare, there is still a lack of empirical benchmarks of standard metrics such as latency, throughput, and synchronization fidelity for contemporary event-streaming platforms [15, 21]. Though wide-ranging reviews on surgical digital twins have highlighted the potential of DTs in pre-, intra-, and postoperative workflows, difficulties remain in harmonizing heterogeneous imaging, kinematics, and physiology under strict latency budgets and in building interoperable, privacy-preserving, and regulatory-compliant DTs in clinical workflows [25]. Thus, while the conceptual capability of EDAs to provide real-time, asynchronous updates to their states is obvious, clear architectural templates are currently absent in the literature on hospital digital twins [4, 14]

The development of high-integration and high-scalability DT architectures also manifests the scalability-privacy duality: Highly actionable operational DT capabilities require perpetual aggregation of hyper-sensitive clinical documents, IoT telemetry, and/or demographic data and are further increased by the ubiquitous use of cloud-based microservice architecture. The attack surface area is large, as evidenced by the 2025 ManageMyHealth data breach in Aotearoa/New Zealand, where over 400,000 highly sensitive medical documents were exfiltrated by threat actors [6]. More recent frameworks propose digital twins and blockchain based on IoT to provide data integrity and immutability, as well as audit trails for decentralized trust without intermediaries [18, 22]. Additionally, systematic reviews of clinical data interoperability models confirm that standards such as HL7 FHIR are a good fit for real-time event-driven data pipelines, and data standards such as DICOM and CDA and data formats such as JSON can be converted to HL7 FHIR to fit almost all clinical use cases [19]. Rapid literature reviews of DTs for healthcare management also identify interoperability, security, and data integration as the most common barriers to common adoption [20]. Future architectures should consider incorporating zero-trust security models and network isolation, as well as interoperable standards (e.g., FHIR-based event schemas) that are scalable , but provide strict patient privacy guarantees [17, 23].

Also related is this trade-off between accuracy and computational efficiency; e.g., studies have shown that tasks such as estimating ED wait times or calculating overtime on ICU beds , lie in solving complex mathematical queuing equations [1]. In order for the digital twin to be fully faithful, additional high-frequency variables such as dynamic nurse-to-patient ratios, variable shift lengths and stochastic patient arrivals are required. While real-time computation and inference on mathematical models across a microservices cluster is computationally intensive; reference architectures for cyber-physical healthcare systems have noted that edge computing is able to reduce the latency of healthcare system solutions since it is able to compute the data closer to its point of generation. However, careful distribution of workloads among edge devices is important since they can easily get overwhelmed [23, 27]. Hospital administrators must balance the need for fine-grained and time-sensitive clinical accuracy against the cost of using cloud resources.

The switch to cloud-native microservices also introduces the risk of vendor lock-in. Many existing systems are commercial off-the-shelf (COTS) products, which have been heavily customized and can have disastrous response times and inflexible upgrade cycles as a result. Hospitals will then try to escape this lock-in through new greenfield microservice implementations, but this may instead result in cloud vendor lock-in, notably through event-streaming and container orchestration services [24]. It is also hard to find the "right service cut" during architectural decomposition, and the effort to refactor poorly-defined boundaries to prevent cascading failures can become an expensive continuing concern [28].

To summarize, for bridging the identified gaps in architecture, we propose a tri-layer conceptual architecture for hospital DT deployments according to referential architectures for cyber-physical healthcare systems [13, 23, 27]:

- Physical Context Layer: The physical hospital environment, which includes IoT vital sign monitors, location-tracking sensors, legacy Electronic Health Record (EHR) databases and human actors, namely patients and clinicians [3] .
- The Event-Driven Microservices Layer: Serves as the asynchronous nervous system. It ingests continuous data from the physical context of the product using message brokers (e.g., Apache Kafka) that publish immutable events. It replaces monolithic bottlenecks with bounded stateless microservices for each business domain (such as admissions, triage, bed management) and uses zero-trust continuous authentication for all inter-service communication [4, 17] .
- The Agent-Based Digital Twin Layer: This is the interactive 'mirror world' where software agents subscribe to the event streams and replicate the hospital digitally in real time. The agents run mathematical capacity models and simulation capabilities, acting as a smart preemptive real-time predictive personal assistant to the healthcare workers before an ambulance arrives (e.g., pre-alerting the trauma team) [3].

Microservices provide the required structural modularity to dynamically scale individual hospital functions that legacy monolithic architectures do not support [16, 24]. Event-driven architectures provide the required data

integration by decoupling data producers from data consumers. Use of semantic interoperability standards (e.g., FHIR) enables interoperability across multiple data sources [19, 21]. While the evidence on healthcare applications is still nascent, and software engineering has used these architectural patterns for decades, the combination of asynchronous event-driven messaging and digital twin architecture means the digital twin layer can receive state updates using the low latencies required for high synchronization fidelity. This turns the digital twin from a 3D representation into a predictive tool that can clear bottlenecks, manage surgical intake, and optimize patient care [25]. .

8. Limitations and future research directions

While this review provides a consolidated framework for understanding event-driven architectures in healthcare, certain scope limitations must be acknowledged. The primary focus remained on the architectural intersection of cloud microservices and digital twins within single-facility operational contexts. Consequently, the review reflects a broader limitation in the current academic literature: empirical evidence of fully deployed, industry-scale event-driven architectures in healthcare remains scarce, with many published studies relying on conceptual models, prototype deployments, or data that lack industry-scale maturity [4]. Furthermore, the selected studies predominantly measure operational throughput metrics, thereby omitting comprehensive analyses of direct clinical impacts.

Furthermore, while reference architectures for cyber-physical healthcare systems have been proposed, empirical validation of event-driven microservice deployments in live hospital settings remains scarce [23, 27].

Addressing these limitations requires a targeted research agenda. Based on the synthesised literature, we identify three high-priority research gaps that must be addressed to advance the field:

Multi-hospital twin federation: Current digital twin deployments are overwhelmingly bounded to single hospitals or specific departments, such as isolated emergency or radiology units [1]. However, patient flow bottlenecks and mass casualty incidents inherently span regional healthcare networks, requiring dynamic patient transfers across various facilities to safely manage critical care resources. Future research must investigate the architectural requirements for "federated" digital twins [29]. This involves establishing secure, cross-organizational event-streaming protocols that allow independent hospital microservices to share real-time capacity and admission data, thereby optimizing regional ambulance routing and resource allocation without violating strict data privacy regulations [30, 31].

AI-augmented predictive twins: While current operational digital twins effectively utilise mathematical queueing models and discrete-event simulations to model backlog and wait times, there is a critical need to transition toward AI-augmented predictive models [1]. The literature demonstrates the potential of machine learning to generate synthetic data and predict clinical disease progression. Future operational research should integrate these advanced artificial intelligence algorithms directly into the event-driven messaging layer [32]. This integration would enable digital twins to evolve from passive mirror worlds into autonomous agents that can proactively predict resource exhaustion and autonomously suggest optimal staffing or scheduling reconfigurations.

Clinical outcome linkage: Finally, the literature currently separates operational digital twins from clinical patient outcomes. Existing studies evaluate architectural success based on reduced patient wait times, lower overtime use, and improved structural packing efficiency [1]. A vital future research direction involves empirically linking these operational efficiencies to concrete clinical health outcomes. Studies must be designed to quantify how event-driven digital twin interventions directly impact critical patient metrics, including inpatient mortality rates, overall lengths of stay, and failure-to-rescue rates [33].

By pursuing this research agenda, the academic and clinical communities can transform hospital digital twins from isolated, descriptive simulations into federated, intelligent, and clinically validated systems that fundamentally improve global healthcare delivery [34].

9. Conclusion

This review set out to address a critical gap in the existing literature regarding the software architectures necessary to support operationally actionable hospital digital twins. While the conceptual appeal of digital twins for optimizing emergency department flow, intensive care unit capacity, and physical hospital layouts is well documented, their practical success relies entirely on real-time synchronization fidelity. Our synthesis establishes that legacy monolithic IT systems are fundamentally ill-equipped to handle the high-frequency, stochastic data streams generated by modern clinical environments. Transitioning to cloud-native microservices provides the requisite scalability and modularity. More crucially, we have demonstrated that integrating event-driven architectures, utilizing asynchronous message brokers, acts as the definitive enabling layer for coupling the physical hospital with its digital replica. This approach minimizes latency, facilitates historical event replay for predictive simulations, and securely decouples disparate data sources, thereby elevating digital twins from static 3D models to dynamic, real-time operational tools.

The findings of this review hold profound implications for hospital informatics practice. As healthcare institutions increasingly adopt digital twins to manage severe resource constraints and time-critical trauma pathways, IT departments must pivot from traditional, synchronous database designs toward asynchronous, stream-processing paradigms. This necessitates a fundamental restructuring of how clinical data are ingested and processed, requiring practitioners to prioritize robust interoperability standards (such as HL7 FHIR) to ensure seamless pipelines between IoT devices, electronic health records, and administrative systems.

From a policy perspective, the architectural shift toward highly distributed, cloud-based digital twins severely exacerbates the tension between operational scalability and patient data privacy. The catastrophic reality of recent healthcare data breaches underscores the urgent need for stringent cybersecurity policies. Healthcare policymakers and hospital administrators must mandate zero-trust security frameworks, strict network isolation, and comprehensive compliance with data protection regulations as prerequisites for deployment. Ultimately, realizing the transformative potential of hospital digital twins requires not just advanced technological adoption but also a corresponding evolution in hospital IT governance and security policy.

Appendix A. PRISMA Search and Screening Protocol

This appendix provides supplementary detail on the systematic search and screening protocol. The search was conducted across PubMed, IEEE Xplore, Scopus, and the ACM Digital Library, bounded to literature published between January 2013 and December 2024.

A.1. Database Search Query

The full Boolean search string applied across all four databases was:

(digital twin OR digital replica) AND (hospital operations OR emergency department OR intensive care unit OR patient flow) AND (microservice OR event-driven architecture OR event streaming OR Kafka OR EventBridge) (A.1)

Following duplicate removal, articles were screened by title and abstract, then assessed by full-text review against defined inclusion and exclusion criteria, documented in accordance with PRISMA guidelines.

Appendix B. Tri-Layer Conceptual Framework

This appendix provides concise definitions for each layer of the tri-layer framework proposed in Section 7.

The Physical Context Layer comprises the real-world hospital environment, including IoT monitors, EHR databases, ADT systems, and human actors whose interactions generate discrete, capturable state-change events.

The Event-Driven Microservices Layer functions as the asynchronous communication backbone, ingesting physical data streams via message brokers and distributing immutable events to bounded, stateless microservices organised around discrete hospital business domains.

The Agent-Based Digital Twin Layer maintains a continuously synchronised digital replica of the physical hospital, executing capacity models, supporting event replay for predictive simulation, and delivering real-time decision support to clinical personnel.

1. **References**

Heydari M, Lai KK, Fan Y, Li X. A review of emergency and disaster management in the process of healthcare operation management for improving hospital surgical intake capacity. Mathematics. 2022 Aug 5;10(15):2784.

2. Singh M, Srivastava R, Fuenmayor E, Kuts V, Qiao Y, Murray N, Devine D. Applications of digital twin across industries: A review. Applied Sciences. 2022 Jun 4;12(11):5727. Croatti A, Gabellini M, Montagna S, Ricci A. On the integration of agents and digital twins in healthcare. Journal of Medical Systems. 2020 Sep;44(9):161.Fritzsch J, Bogner J, Wagner S, Zimmermann A. Microservices migration in industry: Intentions, strategies, and challenges. In 2019 IEEE International Conference on Software Maintenance and Evolution (ICSME) 2019 Sep 29 (pp. 481-490). IEEE. RNZ, Manage My Health reveals scope of data breach. 2026. https://www.rnz.co.nz/news/national/583030/managemyhealth-reveals-scope-of-data-breach
3. Laubenbacher R, Sluka JP, Glazier JA. Using digital twins in viral infection. Science. 2021 Mar 12;371(6534):1105-6.
Björnsson B, Borrebaeck C, Elander N, Gasslander T, Gawel DR, Gustafsson M, Jörnsten R, Lee EJ, Li X, Lilja S, Martínez-Enguita D. Digital twins to personalize medicine. Genome medicine. 2019 Dec 31;12(1):4.
4. Kruse CS, Stein A, Thomas H, Kaur H. The use of electronic health records to support population health: a systematic review of the literature. Journal of medical systems. 2018 Nov;42(11):214. Rudsari HK, Tseng B, Zhu H, Song L, Gu C, Roy A, Irajizad E, Butner J, Long J, Do KA. Digital twins in healthcare: a comprehensive review and future directions. Frontiers in Digital Health. 2025 Nov 18;7:1633539.

5. Fennelly O, Cunningham C, Grogan L, Cronin H, O'Shea C, Roche M, Lawlor F, O'Hare N. Successfully implementing a national electronic health record: a rapid umbrella review. International Journal of Medical Informatics. 2020 Dec 1;144:104281. Finnegan H, Mountford N. 25 years of electronic health record implementation processes: scoping review. Journal of Medical Internet Research. 2025 Mar 3;27:e60077. Tsai CH, Eghdam A, Davoody N, Wright G, Flowerday S, Koch S. Effects of electronic health record implementation and barriers to adoption and use: a scoping review and qualitative analysis of the content. Life. 2020 Dec 4;10(12):327.
6. Roopa MS, Venugopal KR. Digital twins for cyber-physical healthcare systems: architecture, requirements, systematic analysis and future prospects. IEEE Access. 2025 Mar 5.
7. Ferko E, Bucaioni A, Behnam M. Architecting digital twins. Ieee Access. 2022 May 5;10:50335-50.
8. Warrier A. Real-Time Healthcare Event Processing: Stream Analytics for Clinical Decision Support. International Journal of Emerging Research in Engineering and Technology. 2020 Dec 30;1(4):47-54.
9. El Akhdar A, Baidada C, Kartit A, Hanine M, García CO, Lara RG, Ashraf I. Exploring the potential of microservices in internet of things: A systematic review of security and prospects. Sensors. 2024 Oct 21;24(20):6771.
10. Almuseelem WA. Continuous and mutual lightweight authentication for zero-trust architecture-based security framework in cloud-edge computing-based healthcare 4.0. J. Theor. Appl. Inf. Technol. 2024 Jan 15;102(1):1-8.
11. Hemdan EE, El-Shafai W, Sayed A. Integrating digital twins with IoT-based blockchain: concept, architecture, challenges, and future scope. Wireless personal communications. 2023 Aug;131(3):2193-216.
12. Ait Abdelouahid R, Debauche O, Mahmoudi S, Marzak A. Literature review: clinical data interoperability models. Information. 2023 Jun 27;14(7):364.
13. Elkefi S, Asan O. Digital twins for managing health care systems: rapid literature review. Journal of medical Internet research. 2022 Aug 16;24(8):e37641. Krause T, Zickfeld M, Bruchhaus S, Reis T, Bornschlegl MX, Buono P, Kramer M, Mc Kevitt P, Hemmje M. An event-driven architecture for genomics-based diagnostic data processing. Applied Biosciences. 2023 Jun 2;2(2):292-307.
14. Hemdan EE, Sayed A. Smart and secure healthcare with digital twins: a deep dive into blockchain, federated learning, and future innovations. Algorithms. 2025 Jun 30;18(7):401.
15. Rahim M, Lalouani W, Toubal E, Emokpae L. A digital twin-based platform for medical cyber-physical systems. IEEE Access. 2024 Nov 18;12:174591-607.
16. Oumoussa I, Saidi R. The ontology-based mapping of microservice identification approaches: a systematic study of migration strategies from monolithic to microservice architectures. Computers. 2025 Apr 5;14(4):133.
17. Khan AS, Fan F, Kim DD, Alshareef A, Chen D, Kim J, Carter E, Liu B, Rozenblit JW, Zeigler B. A Comprehensive Survey on Surgical Digital Twin. arXiv preprint arXiv:2512.00019. 2025 Oct 28.
18. Katiyar S. Event Driven Architecture for a Health Sciences Customer Using Kafka and Java Microservice. IJLRP-International Journal of Leading Research Publication.;4(12).
19. Roopa MS, Venugopal KR. Digital twins for cyber-physical healthcare systems: architecture, requirements, systematic analysis and future prospects. IEEE Access. 2025 Mar 5.
20. Ratnayake D. Integrated AI-driven marketing growth models for scaling businesses in competitive direct-to-consumer landscapes. IPHO-Journal of Advance Research in Business Management and Accounting. 2025;3(3):01–09.
21. Fauzia A-Clottey P. An integrated operations–leadership framework for enhancing supply chain efficiency and financial oversight. IPHO-Journal of Advance Research in Applied Science. 2025;3(12):42–49.
22. Rubinstein I. Aligning monetization strategy with corporate finance: Performance management in technology-driven advertising businesses. IPHO-Journal of Advance Research in Applied Science. 2025;3(11):01–08.
23. Benneh ND. Institutional governance and risk management in modern financial systems. IPHO-Journal of Advance Research in Business Management and Accounting. 2025;3(12):56–65.
24. Suthari Y, Thakker Z, Govindarajan S, Krishnan N, Raju V, Rao Katta SK. Cost optimization techniques for efficient resource allocation in cloud computing environments. In Proceedings of the 2025 3rd International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI). 2025;793–797. IEEE.
25. Nellutla N. Secure DevSecOps workflows for medical IoT device integration in smart hospitals. International Journal of AI, BigData, Computational and Management Studies. 2022;3(1):114–122.
26. Gollapudi R. Backup integrity and recovery readiness assessment for high-availability databases. Computer Fraud and Security. 2024;23.
27. Gupta, N. Building B2B Advertising Data Intelligence Systems: AI Solutions for Enterprise Marketing, 2025, 10(62s):89-100 https://doi.org/10.52783/jisem.v10i62s.13553