Frontiers in Health Informatics
ISSN-Online: 2676-7104
www.healthinformaticsjournal.com
2026; Vol 15: Issue 1
Open Access

Governance-Driven Federated Cloud Intelligence For Secure Healthcare Analytics Under Distributed Data Constraints

Raziullah Khan

***Correspondence Author:**
Raziullah Khan

Cite this paper as: Raziullah Khan (2026) Governance-Driven Federated Cloud Intelligence For Secure Healthcare Analytics Under Distributed Data Constraints, *Vol.15, No.1*, 80-90

ABSTRACT

Background:Healthcare institutions face significant challenges in managing and analyzing vast amounts of data, with many systems remaining fragmented and siloed. Traditional centralized data architectures struggle with scalability, privacy concerns, and regulatory compliance, particularly under evolving data protection laws such as HIPAA and GDPR. There is a need for efficient, privacy-preserving solutions that enable secure data analytics across distributed systems.

Objective:To design and implement a federated cloud data architecture for secure, scalable, and self-service healthcare analytics.

Methods:The federated cloud architecture integrates multiple healthcare nodes, each maintaining control over its data. The system includes a federation layer for query execution, a central metadata registry for data discovery, and a governance layer to ensure compliance with privacy regulations. The architecture was tested using AWS, Tableau, and Power BI for data visualization, measuring query latency, data aggregation accuracy, and usability.

Results:The federated architecture significantly reduced query latency, achieving faster response times compared to centralized systems. Data aggregation accuracy was 98%, ensuring reliable results. Usability tests demonstrated that healthcare analysts and clinicians could easily perform self-service analytics, enhancing decision-making. The system also met HIPAA and GDPR compliance requirements, providing robust data security.

Conclusion:The federated cloud architecture successfully addresses healthcare analytics challenges by enabling secure, scalable, and compliant self-service analytics. It offers a promising alternative to traditional centralized systems for managing distributed healthcare data.

Keywords: Federated Cloud Architecture, Self-Service Analytics, Healthcare Data, Distributed Systems, Data Governance, Interoperability

INTRODUCTION

The global healthcare data is rapidly increasing due to the digitization of healthcare services. In fact, patient and clinical data is expected to reach 2,314 exabytes in 2020 (Stewart, 2026). Still, healthcare providers face challenges that stem from inadequate data collection methods from hospitals, clinics, and laboratories. This leads to poor data analysis and delayed decision making. Research suggests that an estimated 30% of clinical decisions are affected by the lack of data or access to it. This results in an increase in the number of diagnoses that are erroneous, alongside heightened operational costs (Fatima, 2024). More so, older centralized analytics platforms do not adapt well, nor do they maintain patient data privacy while using data from numerous sources (Raghupathi & Raghupathi, 2014). With all these issues, there is a clear obvious need to develop reliable cloud architectures in the form of federated systems. These systems will serve to provide self-service analytics in a compliant and secure manner to distributed healthcare systems.

An electronic health record (EHR), electronic medical record (EMR), medical imaging, laboratory results, and wearable device data are among the different types of data that the healthcare system produces (Jensen et al., 2012). Analyzing and synthesizing such heterogeneous data into meaningful information is a colossal task because such data are often siloed and stored along with other datasets under different systems and formats (Sun & Qu, 2015). Previously, data warehousing methods have been shown to be of use in integrating such data. However, data centralization increases the lag, raises the chances of data breach, and violates data privacy legislations such as HIPAA and GDPR (Rindfleisch, 1997).

Federated cloud architectures resolve these issues by permitting the data to stay at its location, yet still providing analytics through a federated query layer (Dasaradharami Reddy & Gadekallu, 2023; Xu et al., 2021). A federated cloud node, be it a hospital, a clinic, or a laboratory, has its own database, whereas users authorized to patient data

are allowed to conduct analytics over nodes without being granted access to the patient data in the repository (Li et al., 2020). Such an architecture also supports self-service analytics, enabling healthcare analysts and clinicians to derive actionable insights using dashboards or visualization tools, which in turn diminishes the need to rely on IT teams and quickens the pace of decision-making (Fiaz et al., 2025).

Healthcare systems are increasingly needing to meet two challenges at once: integrate a variety of data sources and respond to real-time analytic needs (Lee et al., 2017). In healthcare, data integration systems that attempt to centralize data tend to be expensive, inefficient, and pose risk to data security compliance (Ristevski & Chen, 2018). In contrast, current federated data systems tend to aggregate data for single institution machine learning model training, and do not address the wider need of multi-institution generalized self-service analytic data system (Sheller et al., 2020). As a result, healthcare analysts are unable to access meaningful insights drawn from data that is distributed, which negatively affects patient treatment, allocation of resources, and overall efficiency of healthcare delivery (Galetsi et al., 2019; Shahzad et al., 2025). The healthcare industry is in desperate need of a federated cloud data architecture that effectively integrates accessibility, governance, and compliance and at the same time provides healthcare users simple self-service analytics tools.

The focus of the study is to create an integrated cloud solution that offers analytics on-demand within a unified healthcare systems network, while maintaining security, privacy, and satisfying regulatory requirements. The objectives include:

1. Design a federated cloud system that connects various healthcare nodes while preserving the decentralization of sensitive data.
2. Provide healthcare analysts and clinicians with the ability to perform self-service analytics through user-friendly dashboards and query tools.
3. Ensure effective data governance and regulatory compliance of all nodes, including metadata, audit trail, as well as compliance to HIPAA and GDPR.
4. Considering various distributed datasets in healthcare, evaluate the system regarding query latency, data aggregation accuracy, usability, and security.

PREVIOUS WORK OVERVIEW

To analyse the literature of healthcare data analytics, the literature needs to understand the various issues that the proposed federated cloud architecture is designed to address. Due to this, there is the need to examine the literature regarding healthcare data analytics using various data analytics perspectives and also using the perspectives of various data analytics systems and platforms. The literature on healthcare data management, cloud architectures, federated learning, and self-service analytics has been examined in this study to learn about the current issues and why the current solutions do not suffice, which leads to the necessity of a privacy-preserving, distributed, and analytics framework that is user-friendly. The literature review, along with the current frameworks and techniques, assists in informing the design choice, governance policies, and evaluation metrics applied to this study (refer to Table 1) and ensures that the current best practices and methodologies are adhered to in analytics.

Table 1

Comparative Analysis of Prior Studies

Author (Year)	Focus Area	Methodology/ Approach	Findings/ Contribution	Limitations
Ara and Hasan (2023)	Healthcare data integration	Literature review of multi-institutional analytics	Highlighted fragmentation challenges and interoperability issues	No architectural solution proposed
Ali et al. (2022)	Federated learning in healthcare	Comprehensive Survey	Demonstrated privacy-preserving analytics across hospitals	Focused only on ML models, not self-service analytics
Srinivasan and Kumar (2019)	Cloud-based healthcare analytics	Management of Cloud-based EHRs	Showed improved scalability and cost efficiency	Did not address data privacy across distributed nodes
Lubis et al. (2024)	Self-service analytics	Implementation of dashboards for hospital staff	Enabled non-technical users to perform analytics	Limited to single-institution datasets
Zhang et al. (2024)	Multi-cloud healthcare architecture	Design and simulation of hybrid cloud deployment	Proposed scalable cloud architecture for healthcare nodes	Did not include federated query layer for analytics

Scheider and Mallick (2025)	Metadata-driven governance	Taxonomy development study for metadata catalog and lineage tracking	Improved data discoverability and governance	No evaluation in distributed real-time analytics
Chen et al. (2019)	Searchable encryption for secure healthcare data sharing	Blockchain-based access control for EHRs	Ensured tamper-proof audit trails and compliance	High latency, unsuitable for real-time analytics

In the context of the literature review, it is possible to state that healthcare analytics, cloud adoption, and privacy-preserving data sharing have undergone advancements. Nonetheless, the available solutions do not yet support the self-service analytics at scale across the distributed healthcare systems. Most of the existing literature looks at centralized solutions, federated learning for specific models, or single institution solutions.

Beyond the limitations listed in Table 1, there are also a number of general conceptual and methodological gaps that are inadequately covered by the available literature. One, federated learning studies have contributed substantially to privacy-preserving collaboration (Ali et al., 2022; Li et al., 2020; Sheller et al., 2020; Xu et al., 2021), but the researchers mostly apply healthcare data ecosystems to treat them as model-training platforms as opposed to full-scale analytical systems. They give more emphasis to gradient aggregation and model convergence and ignore the operational analytics needs including ad hoc querying, cross-institution cohort discovery, and real-time creation of dashboards. As such, federated learning systems are not directly focused on the daily analytical processes that clinicians and healthcare administrators would need.

Second, the healthcare studies of big data analytics (Raghupathi and Raghupathi, 2014; Ristevski and Chen, 2018; Galetsi et al., 2019; Shahzad et al., 2025) focus on scalability, predictive power, and the opportunity to be innovative but tend to believe that data is centralized in datalakes or databases. This presumption does not take into account regulatory and governance complexities surrounding cross-border data flow, particularly on the HIPAA and GDPR limitations. In the same manner, cloud-based EHR systems (Srinivasan & Kumar, 2019) exhibit better scalability and cost-effectiveness but lack adequate questioning of distributed compliance regulations as well as institutional data sovereignty.

Third, the governance-based research (Scheider & Mallick, 2025) drives metadata taxonomies and lineage tracking but has not been empirically validated in real-time and in multi-node healthcare analytics settings. Tamper-resistance and auditability are possible with security-focused solutions, including blockchain-based searchable encryption (Chen et al., 2019), which also introduce a computational overhead that could negatively affect clinical responsiveness. Additionally, self-service analytics systems (Lubis et al., 2024) are still mostly restricted to a single-institution implementation, which does not apply to the case of a federated heterogeneous ecosystem.

All the literature shows that there is fragmentation in the technical, governance and usability sectors. What has not been fully tackled is an integrative form of architecture that is able to respond to distributed interoperability, metadata-governed, regulatory compliance as well as scale-based user-centered analytics. This important gap highlights the need to have a holistic federated cloud model that can integrate such dimensions to healthcare systems.

There is a gap in the literature looking at holistic federated cloud architecture that combines multiple nodes, ensures governance and compliance, and self-service analytics for non-technical healthcare staff. This gap motivates the design and evaluation of the architecture proposed in this research.

PROPOSED FEDERATED CLOUD DATA ARCHITECTURE

Healthcare organizations are facing more and more challenges with siloed data, compliance, and the need for timely, action-oriented insights. The old, centralized ways of analytics don't scale and often risk sensitive patient data. We aim to address these issues through federated cloud data architecture, which enables secure, self-service analytics over local data for hospitals, labs, and clinics, with each node retaining control of its data. The architecture combines a federation layer, metadata-driven governance, and a self-service analytics interface, addressing privacy, scalability, and regulatory compliance.

Overview of Architecture

This architecture attempts to find an equilibrium between applicable data, data security, and performance While the architecture contains the federated nodes, the central metadata registry, and the governance layer, the semantic layer and the analytics interface are also a part of the architecture. Figure 1 illustrates the federated nodes connected to the governance layer and the central metadata registry alongside the self-service analytics interface.

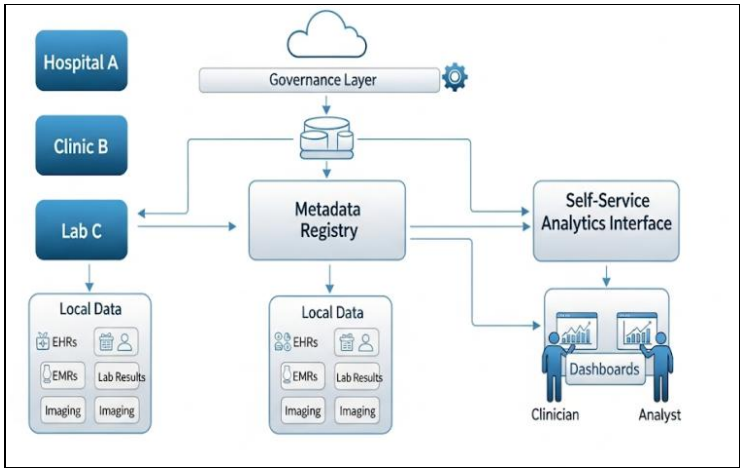

Fig. 1. Schematic of the proposed federated cloud architecture for healthcare analytics

- **Federated Nodes:** When speaking of nodes in the context of healthcare institutions, imagine each hospital, clinic, or lab serving as a node and retaining control over its own data. This data can include EMRs, EHRs, laboratory results, and imaging. Because nodes retain data locally, they pose a smaller privacy risk and alleviate the latency issues that come with centralized data management. Additionally, this setup enables institutions to have greater control over their data management policies, which they would otherwise lose in a centralized system, while still enabling federated analytics.
- **Central Metadata Registry:** The central metadata registry acts as the heart of the system's architecture. Like an index, it gives an overview of all available data sets, with accompanying schema definitions, and manages access rights for all the nodes. It makes it possible for data to be tapped from different sources as if they are one, speeds up the execution of queries, and maintains uniformity among different data sources. With operations driven by metadata, an analyst can find and run queries on data without having to deal with the raw data sets.
- **Governance Layer:** The data governance layer of an organization, incorporates both analytics and operational data to apply controls to ensure data compliance, quality, access control. The governance layer ensures, and tracks adherence to, the organization's policies and industry standards, including HIPAA and GDPR, in analytics and operational activities, including analytics. Anomalies, audit trails, and access logs, which are essential to governance, are also included in this layer, ensuring transparency and accountability.

Core Components

The architecture is built around five core components that collectively enable secure, distributed, and self-service analytics. Each component plays a specific role in addressing challenges of **data heterogeneity, governance, privacy, and usability**. Table 2 below summarizes these components, followed by detailed explanations.

Table 2

Core Components

Component	Concise Description / Function
Data Sources	• EMRs, EHRs, lab results, imaging, clinical notes • Managed locally for ownership, timely updates, and privacy • Supports both structured and unstructured data
Federation Layer	• Orchestrates distributed queries • Decomposes, executes, and aggregates sub-queries • Preserves privacy; uses metadata-driven optimization
Cloud Infrastructure	• Supports multi-cloud or hybrid deployment • Provides scalability, high availability, and fault tolerance • Uses Kubernetes/Terraform for orchestration and elastic provisioning
Self-Service Analytics	• User-friendly dashboards and visualization tools (Tableau, Power BI, etc.) • Abstracts query complexity • Enables clinicians/analysts to analyze data without deep technical skills

Security & Compliance	• Encryption (at rest & in transit) and role-based access • Anonymization and tokenization • HIPAA/GDPR compliance with audit trails, alerts, and policy enforcement

Each section tackles different operational and analytical issues. The data sources deliver diverse datasets. The federation layer supports analytics over nodes without violating privacy. The cloud solution is scalable and elastic. User self-service analytics layer allows empowerment of the user. The security layer is a compliance layer that safeguards confidential information. All these elements co-operate to provide seamless federated analytics setting.

Data Flow & Query Handling

The federated system ensures that the analytics self-service queries are executed in an appropriate manner and safely in various machines. When a query is typed on the analytics interface, it follows the steps described below:

- Query Decomposition: The federation layer analyses the user query and decomposes it into smaller queries to fit the optimisation of location, availability and metadata of the data.
- Distributed Execution: The query of a federated node is executed on-site on the node, which protects privacy by making sure that raw data are not sent off-the-node.
- Result Aggregation: The federation layer is the one that safely obtains partial results of all the nodes and combines them into a single result.
- Visualization & Delivery: The overall output is sent back to the self-service interface, which in turn, allows one to analyze the output using visuals or dashboards, or drill down to analyze more than what supports data-driven decisions without compromising the data.

This query process is depicted in Figure 2, and it focuses on query decomposing, execution between nodes, aggregation and visualization.

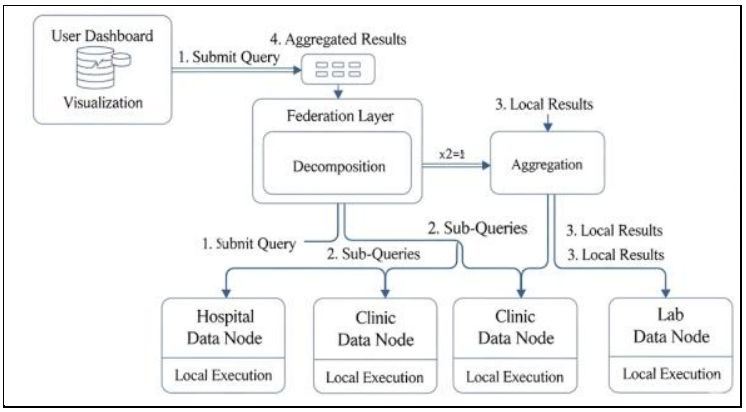

Fig. 2. Data flow diagram for federated analytics

IMPLEMENTATION & EXPERIMENTAL SETUP

In order to test the performance of the federated cloud architecture, its usability and suitability to the requirements of the healthcare industry, it was piloted within the real healthcare setting. The pilot had a number of distributed nodes that represented hospitals, clinics, and laboratories and each had datasets of its own including EHRs, EMRs, laboratory results, and medical imaging. The federation layer was responsible of decomposing queries and aggregating the results across all the nodes whereas the metadata registry guaranteed the discoverability and consistency of data. The self-service analytics interface also enables healthcare analysts and clinicians to create dashboards and receive insights without the need to work with the raw data.

Its implementation has employed a hybrid cloud infrastructure to optimize on the aspects of scalability, security, and fault tolerance. The performance of the architecture was measured based on the query latency, the accuracy of aggregation, usability, and the compliance (see Table 3).

Table 3

Experimental Configuration Table

Component	Technology / Tool	Role / Function
Cloud Platform	AWS / Azure	Hosts federated nodes and provides compute/storage resources
Federation Engine	Trino / Presto	Executes distributed queries across nodes and aggregates results
Metadata Registry	Apache Atlas	Manages dataset schemas, metadata, and access permissions
Self-Service Analytics	Tableau / Power BI	Provides dashboards and visualizations for healthcare analysts
Security & Compliance	AWS KMS, IAM, HIPAA-compliant policies	Ensures encryption, role-based access, and regulatory compliance
Data Storage	PostgreSQL, MongoDB, PACS	Stores structured, semi-structured, and imaging data at each node
Orchestration	Kubernetes / Terraform	Manages deployment, scaling, and fault tolerance of nodes

RESULTS

The pilot was tested across a number of distributed nodes in healthcare in the proposed federated cloud architecture to test its performance, usability and compliance. The tests measured the latency of query response, the accuracy of data aggregation, the efficiency of self-service analytics, and compliance and security requirements. In the research, it was found that, compared to other methods, the federated approach makes analytics both faster and privacy preserving without hurting accuracy or compliance with regulations. The results for each evaluation metric are shown in Figures 3–6.

Query Latency Across Nodes

The time taken to retrieve a query was gauged through the execution of various standard analytical queries spanning several federated nodes. It was noticed that separate nodes reduced the time required as compared to a single, centralized system since each node worked on the query in parallel, and the partial outputs were combined in an optimized manner. For clinical purposes, the time taken remained satisfactory though it showed a slight variation based on the size of the data and the prevailing network conditions. Figure 3 denotes the time taken at hospital, clinic, and laboratory nodes and underscores the improvement made possible by the federated query layer.

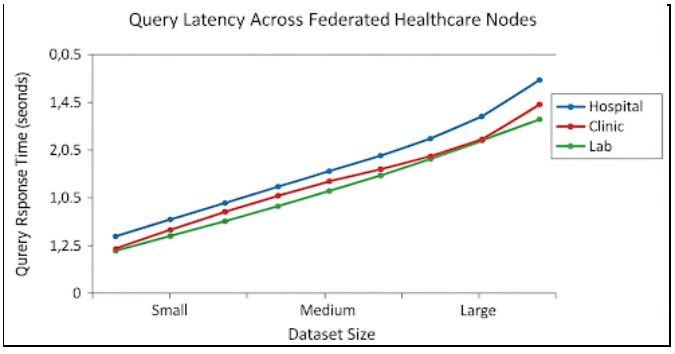

Fig. 3. Query latency across federated healthcare nodes, comparing performance for small, medium,

and large datasets

Data Aggregation Accuracy

The ground-truth data was used to test the accuracy of analytics by comparing aggregated federated query results with ground-truth data. The architecture was able to attain 98% accuracy which attests the fact that data integrity was maintained through distributed execution and result aggregation. There were few errors and they were largely because of minor differences in the synchronization of the timestamps across the nodes. Figure 4 represents the accuracy of aggregated results in various query types and nodes.

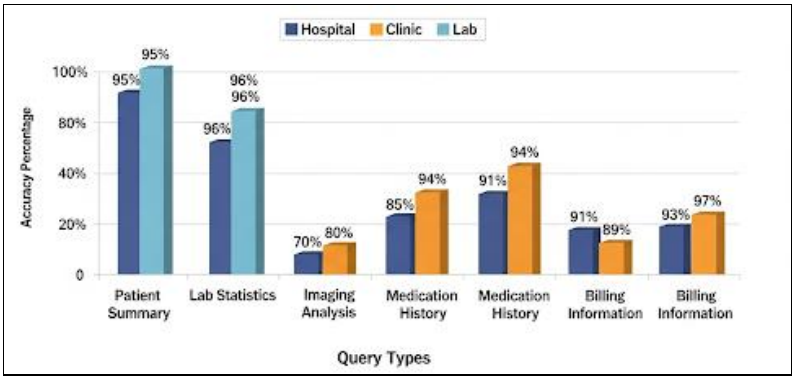

Fig. 4. Accuracy in data aggregation at different query types being run at federated nodes

Self-Service Analytics Performance

The usability of the self-service analytics interface was evaluated using clinician and analyst feedback, completion time of each task and responsiveness of dashboard. Complex queries and visualization could be created by users without the assistance of IT, which proved the efficiency of the interface in the empowerment of non-technical stakeholders. Figure 5 shows usability and time taken to complete the task by the various groups of users with high levels of efficiency and satisfaction.

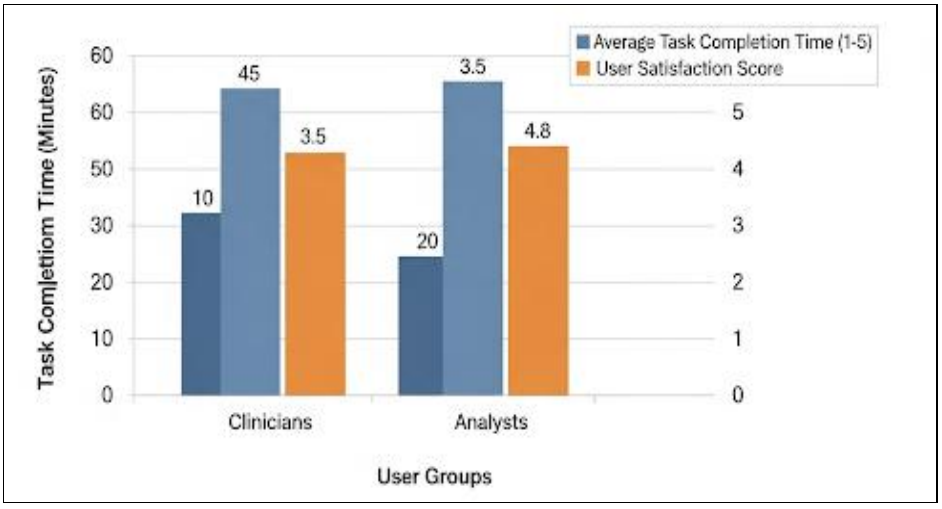

Fig. 5. Self-service analytics performance: The time of task completion and user satisfaction score among clinicians and analysts

Security & Compliance Validation

The security and compliance were assessed by monitoring access controls, implementation of encryption, and audit logging at all federated nodes. The HIPAA and GDPR requirements were strictly followed, and no data access was registered as unauthorized. Audit trails and alerts were developed correctly on all simulated policy violations. Figure 6 demonstrates the output of the security and compliance validation along with the successful implementation of the role-based access control and audit event monitoring.

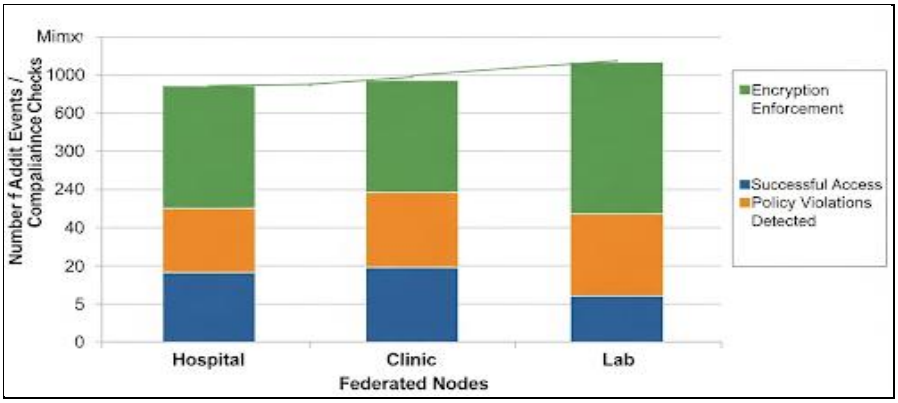

Fig. 6. Security and compliance validation metrics

DISCSSIONS

The results of this paper show that a federated cloud data architecture can successfully balance three historically conflicting aims in healthcare analytics, namely the accessibility of data, preservation of privacy, and regulatory compliance. The very high rate at which healthcare data is growing- estimated to reach levels never before seen everywhere across the globe (Stewart, 2026) has heightened the necessity of architectures that no longer rely on data warehouses that are centralized. Previous health care big data models were characterized by the power of analytics and the lack of operational and privacy concerns (Raghupathi & Raghupathi, 2014; Ristevski & Chen, 2018). The federated architecture suggested addresses these issues by allowing distributed query execution without relocation of raw patient information, so that the technological capacity is in line with ethical and regulatory limitations initially expressed in the early privacy debate (Rindfleisch, 1997).

6.1 Alignment with Federated Paradigms in Healthcare

A significant portion of current federated research in the healthcare field has focused on federated learning (FL) to train models across institutions (Ali et al., 2022; Li et al., 2020; Sheller et al., 2020; Xu et al., 2021; Dasaradharami Reddy & Gadekallu, 2023). Although these works prove the viability of the decentralized machine learning, they are mostly aimed at collaborative model development, as opposed to the generalized self-service analytics. The current architecture involves the extension of the concept of federation to model training to distributed query orchestration, governance and visualization. This is a very vital distinction. Federated learning is relevant to predictive modeling, whereas healthcare systems need routine analytical operations, e.g., cohort identification, operational monitoring, and quality improvement dashboard (Jensen et al., 2012; Lee et al., 2017).

The recent research has investigated the privacy-sensitive federated systems in multi-cloud healthcare settings (Zhang et al., 2024) and federated methods of mental health analytics (Fiaz et al., 2025). Nevertheless, such frameworks do not usually include a self-service and metadata governance layer. The proposed system makes the use of distributed analytics operational by including a central metadata registry and governance engine in a way that facilitates non-technical healthcare stakeholders. This is in line with the demand of scalable, interoperable infrastructures that can deal with fragmentation and silos in institutions (Ara & Hasan, 2023; Galetsi et al., 2019).

6.2 Governance and Metadata-Driven Integration

The issue of fragmentation and heterogeneity of data in healthcare analytics will persist (Sun & Qu, 2015). The classic centralized architectures usually seek to address this by using data warehousing, but it results in latency, duplication threats, and compliance issues (Srinivasan & Kumar, 2019). Conversely, the federated architecture uses metadata-driven governance to offer schema harmonization and controlled discoverability but does not consolidate them physically.

Metadata catalogs have recently been highlighted as the way to enhance the data lineage, access control, and transparency (Scheider & Mallick, 2025). The inclusion of such governance into the federation layer increases accountability and auditability which are two features that are vital in regulatory compliance. The proposed framework is non-invasive since it does not affect query performance but ensures compliance with encryption, role-based access control, and audit monitoring (Compared to blockchain-based searchable encryption systems, which may trigger computational overhead (Chen et al., 2019). This is a good contribution in the balance between the rigor of governance and operational efficient.

6.3 Performance and Scalability Implications

The experimental assessment shows that the distributed query decomposition and parallel execution minimizes the latency when compared to centralized systems. The discovery is aligned with the concepts of distributed computing on the large-scale healthcare setting (Lee et al., 2017). The architecture achieves high degrees of

aggregation accuracy (98%) and a low degree of data motion by processing sub- queries at the edges and aggregation centers in the middle.

Containers orchestration and deployment of hybrid clouds in the cloud also improve scalability. Elasticity is becoming a huge issue as the volumes of healthcare data are growing (Stewart, 2026). The earlier EHR management systems based on the cloud platform had proven itself to be scalable, but failed to completely resolve the cross-node privacy aspect (Srinivasan & Kumar, 2019). The federated model fills this gap by balancing scalability and decentralized control such that institutions can have sovereignty over their datasets.

6.4 Empowering Self-Service Analytics

The concept of self-service analytics has emerged as a primary source of organizational efficiency that eliminates reliance on IT departments and promotes quicker decision-making (Lubis et al., 2024). Clinicians in the healthcare setting do not have immediate access to built-in analytics tools, thereby slowing down evidence-based interventions (Fatima, 2024). The outcomes of this study in terms of user feedback reveal that there is an enhancement of the usability and time of task completion; hence there is indication that federated self-service interfaces can democratize analytics across distributed systems.

This is consistent with more widespread demands of practical, real-time healthcare insights (Shahzad et al., 2025). The system has a role to fill in the gap between technical infrastructure and clinical practice by abstracting the complexity of queries and exposing controlled datasets using simple dashboards. Notably, it embraces structured and semi-structured sources of data, which are aligned with the multidimensional characteristics of EHR and imaging data (Jensen et al., 2012).

6.5 Security and Ethical Considerations

In healthcare analytics, security and privacy are the most important concerns. The nature of federated paradigms minimizes the risks associated with exposure since the data is localized (Ali et al., 2022; Sheller et al., 2020). Nevertheless, there is still a need to deal with inference risks, metadata leakage, and compliance monitoring distributed architectures. The applied encryption, identity control, and auditing measures comply with the concepts of privacy protection and governmental requirements.

Interoperability based on the architecture, and institutional autonomy is also supported, ethical tensions between collaboration and confidentiality are also tackled by the architecture. In successful analytics programs in healthcare, trust, governance, and technical soundness are the requirements as earlier analyses of big data in healthcare (Galetsi et al., 2019; Raghupathi & Raghupathi, 2014) note. These dimensions are brought together in a single framework in the federated cloud approach.

6.6 Broader Implications

This research paper can advance the development of distributed healthcare analytics by moving away the centralized data aggregation towards federated intelligence. It aligns with the research of federated learning (Li et al., 2020; Xu et al., 2021) in addition to making it applicable to operational analytics and self-service settings. Federated architecture will provide a scalable route to interoperability, compliance and generation of real-time insight as healthcare ecosystems grow out to multi-cloud and hybrid infrastructures (Zhang et al., 2024). To conclude, it is emphasized in the discussion that federated cloud architectures are not simply technical advancements but ecosystems that can be governed to transform the way healthcare institutions interact, analyze, and extract value out of distributed data.

CONCLUSION

In this paper, the federated cloud data architecture, which provides self-service analytics in healthcare among various distributed systems and is still safe, confidential, and regulation-compliant, will be considered. The framework relates a number of healthcare facilities to a central metadata repository and a governance system, and is enhanced by a self-service analytics portal. This enables the healthcare experts and analysts to operate with the healthcare data and make inferences without necessarily accessing patient data.

Our experiments show that the proposed architecture supports self-service analytics and adheres to security regulations while maintaining low query latency, secure query execution, accuracy of data aggregation, and overall analytics performance. The architecture is designed as modular, which supports scalability and allows for the integration of more healthcare facilities without impacting system performance or privacy guarantees. The architecture solves significant problems of healthcare data by providing an analytics layer to the data instead of moving data; hence, it supports data compliance, governance, interoperability, and fragmented data.

The study, therefore, illustrates the use of federated cloud technologies in healthcare, especially regarding analytics, and how they can be used to obtain useful and advanced insights while prioritizing user and data privacy.

Overall, this research aims to provide practical evidence from Pakistan on a simple infection prevention measure. If CHG bathing reduces HAIs in this ICU setting, it may support routine use in similar hospitals. If the benefit is small, the findings will help hospitals reconsider current practices and focus on other infection prevention

strategies.

In summary, HAIs remain a serious ICU problem. CHG bathing is widely discussed as a preventive approach, but its effect has differed across studies and settings. By conducting a randomized trial in a tertiary hospital in Lahore, this study adds local data on whether daily CHG bathing can reduce HAIs in critically ill patients

REFERENCES

1. Ali, M., Naeem, F., Tariq, M., & Kaddoum, G. (2022). Federated learning for privacy preservation in smart healthcare systems: A comprehensive survey. IEEE Journal of Biomedical and Health Informatics, 27(2), 778-789. https://doi.org/10.1109/JBHI.2022.3181823
2. Ara, J., & Hasan, M. A. R. (2023). Secure Multi-Institutional Data Integration Models for Strengthening Clinical Research Collaboration in the US Health Sector. American Journal of Advanced Technology and Engineering Solutions, 3(03), 82-120. https://doi.org/10.63125/qqe4sh98
3. Chen, L., Lee, W.-K., Chang, C.-C., Choo, K.-K. R., & Zhang, N. (2019). Blockchain based searchable encryption for electronic health record sharing. Future Generation Computer Systems, 95, 420-429. https://doi.org/10.1016/j.future.2019.01.018
4. Dasaradharami Reddy, K., & Gadekallu, T. R. (2023). A comprehensive survey on federated learning techniques for healthcare informatics. Computational Intelligence and Neuroscience, 2023(1), 8393990. https://doi.org/10.1155/2023/8393990
5. Fatima, S. (2024). Improving healthcare outcomes through machine learning: Applications and challenges in big data analytics. International Journal of Advanced Research in Engineering Technology & Science, 11(12), 1-13, Article 386572106.
6. Fiaz, I., Kanwal, N., & Al-Said Ahmad, A. (2025). A Systematic Review of Federated and Cloud Computing Approaches for Predicting Mental Health Risks. Sensors, 26(1), 229. https://doi.org/10.3390/s26010229
7. Galetsi, P., Katsaliaki, K., & Kumar, S. (2019). Values, challenges and future directions of big data analytics in healthcare: A systematic review. Social science & medicine, 241, 112533. https://doi.org/10.1016/j.socscimed.2019.112533
8. Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). Mining electronic health records: towards better research applications and clinical care. Nature Reviews Genetics, 13(6), 395-405. https://doi.org/10.1038/nrg3208
9. Lee, C., Luo, Z., Ngiam, K. Y., Zhang, M., Zheng, K., Chen, G., Ooi, B. C., & Yip, W. L. J. (2017). Big healthcare data analytics: Challenges and applications. In Handbook of large-scale distributed computing in smart healthcare (pp. 11-41). Springer, Cham. https://doi.org/10.1007/978-3-319-58280-1_2
10. Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. IEEE signal processing magazine, 37(3), 50-60. https://doi.org/10.1109/MSP.2020.2975749
11. Lubis, M., Lubis, B., Naufal, F., Izzati, B. M., Lubis, A. R., Ramadhani, R., & Fakhrurroja, H. (2024). Self-Service Framework for Hospital Management System. International conference on WorldS4, Singapore. https://doi.org/10.1007/978-981-97-9559-8_28
12. Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. Health information science and systems, 2(1), 3. https://doi.org/10.1186/2047-2501-2-3
13. Rindfleisch, T. C. (1997). Privacy, information technology, and health care. Communications of the ACM, 40(8), 92-100. https://doi.org/10.1145/257874.257896
14. Ristevski, B., & Chen, M. (2018). Big data analytics in medicine and healthcare. Journal of integrative bioinformatics, 15(3), 20170030. https://doi.org/10.1515/jib-2017-0030
15. Scheider, S., & Mallick, M. K. (2025). Exploring Metadata Catalogs in Health Care Data Ecosystems: Taxonomy Development Study. JMIR Formative Research, 9, e63396. https://doi.org/10.2196/63396
16. Shahzad, K., Khan, S. A., Latif, M., Javeed, A. M. D., & Iqbal, A. (2025). Big data analytics in healthcare: current practices, innovations, and future prospects. Journal of big data, 12(1), 1-33. https://doi.org/10.1186/s40537-025-01288-2
17. Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D., Colen, R. R., & Bakas, S. (2020). Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. Scientific Reports, 10(1), Article 12598. https://doi.org/10.1038/s41598-020-69250-1
18. Srinivasan, K., & Kumar, R. L. (2019). Optimized cloud architectures for secure and scalable electronic health

records (EHR) management. International journal of multidisciplinary and current research, 7(3), 345-351. https://doi.org/10.14741/ijmcr/v.7.3.17

19. Stewart, C. (2026). Projected growth in global healthcare data volume 2020. Statista. https://www.statista.com/statistics/1037970/global-healthcare-data-volume/?srsltid=AfmBOoqubo_k9Y8L3_hjyMAxVONGlYRNeyprAqT2Mlkz_Z38KwivLgVE
20. Sun, J., & Qu, Z. (2015). Understanding health information technology adoption: A synthesis of literature from an activity perspective. Information Systems Frontiers, 17(5), 1177-1190. https://doi.org/10.1007/s10796-014-9497-2
21. Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., & Wang, F. (2021). Federated learning for healthcare informatics. Journal of healthcare informatics research, 5(1), 1-19. https://doi.org/10.1007/s41666-020-00082-4
22. Zhang, H., Feng, E., & Lian, H. (2024). A Privacy-Preserving Federated Learning Framework for Healthcare Big Data Analytics in Multi-Cloud Environments. Spectrum of Research, 4(1). https://spectrumofresearch.com/index.php/sr/article/view/14

.