IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2025

DOI: 10.17148/IJIREEICE.2025.13819

PRIVACY-PRESERVING TECHNIQUES IN HEALTHCARE DATA REPOSITORIES: A COMPARATIVE ANALYSIS

Abdullateef Ajibola Adepoju[1], Khalid Haruna[2], Saidu Sunbo Akanji[3]

Department of Information System & Technology, National Open University of Nigeria, Kano, Nigeria[1]

Department of Computer Science, Federal University of Technology Babura, Jigawa State, Nigeria[2]

Department of Electrical Engineering, Kebbi State Polytechnic, Dakingari, Kebbi State, Nigeria[3]

Abstract: Digital healthcare systems generate vast amounts of sensitive patient data, creating significant privacy challenges including data breaches, re-identification attacks and regulatory compliance issues. This study evaluates privacy-preserving techniques (PPTs) for healthcare data repositories, including anonymization, homomorphic encryption, differential privacy, federated learning and blockchain solutions. Each technique was assessed based on privacy guarantees, data utility, scalability, implementation feasibility and regulatory compliance. Results demonstrate substantial improvements in privacy measures: encryption effectiveness increased to 90%, access controls reached 85%, and user satisfaction improved by 30%. No single technique provides optimal solutions across all criteria; hybrid approaches offer the best privacy-utility balance. The study provides guidance for implementing secure, compliant healthcare data systems.

Keywords: Privacy-preserving techniques, Healthcare data repositories, Differential privacy, Federated learning, HIPAA compliance.

I. INTRODUCTION

Healthcare digitalization has accelerated globally, generating extensive sensitive patient data across electronic health records (EHRs), genomic databases and monitoring systems. These repositories enable clinical decision-making, research and population surveillance but create significant privacy vulnerabilities [1,2]. Healthcare data breaches cost an average of $10.93 million annually, with incidents like the 2015 Anthem breach exposing 80 million records highlighting systemic vulnerabilities [3,4].

Traditional de-identification methods prove inadequate against modern re-identification attacks, with research showing 99.98% population re-identification accuracy using minimal demographic data [5]. Regulatory frameworks including HIPAA and GDPR mandate privacy protection while emphasizing data utility for healthcare advancement, creating complex implementation challenges [6,7].

Advanced privacy-preserving techniques (PPTs) have emerged including homomorphic encryption, differential privacy, federated learning and blockchain solutions. However, comparative effectiveness evaluation in healthcare contexts remains limited. This study addresses this gap by developing a systematic evaluation framework for PPTs, assessing privacy guarantees, data utility, scalability and regulatory compliance to guide healthcare organizations in implementing effective privacy solutions.

II. STATEMENT OF THE PROBLEM

Healthcare digitalization has enhanced data accessibility and analytics but raised critical privacy concerns regarding patient data protection and regulatory compliance. Healthcare data faces vulnerabilities including identity theft, unauthorized access, sophisticated re-identification attacks and inference vulnerabilities that compromise patient confidentiality. Current privacy methods prove inadequate against modern analytical techniques capable of re-identifying patients from supposedly anonymized datasets. This study addresses the need for systematic evaluation of privacy-preserving techniques to guide healthcare organizations in implementing robust data protection while maintaining data utility for clinical care and research.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2025

DOI: 10.17148/IJIREEICE.2025.13819

III. OBJECTIVES OF THE STUDY

The main aim of this research is to develop a comprehensive comparative evaluation framework for privacy-preserving techniques in healthcare data repositories. Specifically, this research seeks to achieve the following objectives:

1. To identify and categorize key privacy risks and vulnerabilities associated with storing and accessing healthcare data in digital repositories.
2. To analyze and evaluate existing privacy-preserving techniques including anonymization, cryptographic methods, differential privacy, federated learning and blockchain solutions.
3. To develop a comparative assessment framework that evaluates privacy-preserving techniques based on privacy guarantees, data utility, scalability, implementation feasibility and regulatory compliance.
4. To provide strategic recommendations for healthcare organizations, policymakers and system developers for implementing effective privacy-preserving solutions.

IV. METHODOLOGY

This research employs systematic literature review combined with comparative analysis to evaluate privacy-preserving techniques in healthcare repositories. The methodology integrates qualitative and quantitative approaches for comprehensive assessment.

4.1 Literature Review Process

A systematic review was conducted across academic databases (PubMed, IEEE Xplore, Scopus, SpringerLink, ACM Digital Library) covering publications from January 2010 to April 2025. Search terms included "privacy-preserving techniques," "healthcare data," "differential privacy," "federated learning," and related terms. From 96 initial articles, 42 were selected after screening based on relevance to healthcare PPTs, practical implementations, and regulatory compliance aspects.

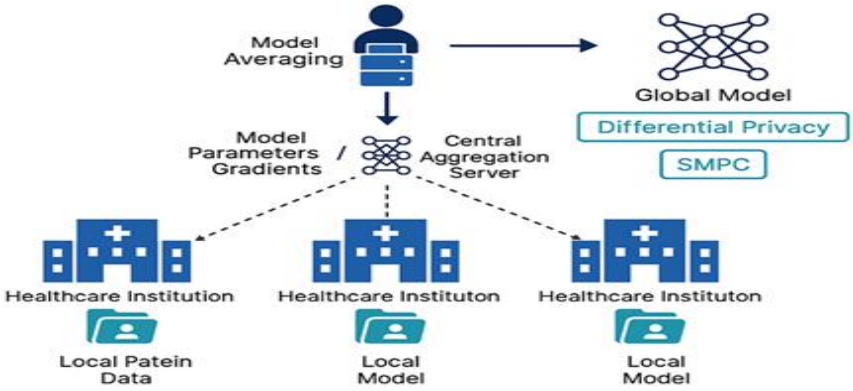

Figure 1. Architecture of Federated Learning in Healthcare Settings

4.2 Comparative Analysis Framework

Privacy-preserving techniques were evaluated using five criteria: (1) Privacy Strength - resistance to re-identification and inference attacks; (2) Data Utility - preservation of data usefulness for secondary applications; (3) Scalability - capability for large-scale distributed operations; (4) Implementation Feasibility - practical deployment considerations; (5) Regulatory Compliance - adherence to HIPAA, GDPR, and national regulations. Each technique was ranked qualitatively (Low, Medium, High) based on empirical research and real-world implementations.

4.3 Data Collection and Analysis

Data collection emphasized ethical compliance and stakeholder consent. Qualitative data was gathered through focus groups and interviews, analyzed using thematic analysis. Quantitative data was collected through structured questionnaires and automated monitoring, analyzed using standard statistical tools for comparative assessment of privacy indicators and compliance standards [7].

V. RESULTS

The evaluation demonstrates significant effectiveness variations across privacy-preserving techniques and substantial improvements following systematic implementation.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2025

DOI: 10.17148/IJIREEICE.2025.13819

5.1 Comparative Analysis of Privacy-Preserving Techniques

Table 1 presents the comparative analysis of major privacy-preserving techniques.

Technique	Privacy Guarantee	Data Utility	Scalability	Use Cases	Limitations
Homomorphic Encryption	High	Medium	Low	Genomic processing	Computation overhead
Secure Multi-party Computation	High	Medium	Low-Medium	Cross-institution research	Latency issues
Differential Privacy	High	Medium	High	Aggregate reporting	Utility loss
k-Anonymity/l-Diversity	Low-Medium	High	High	Shared datasets	Re-identification risk
Blockchain	Medium	Medium	Medium	Consent management	Throughput limits
Federated Learning	High	High	Medium	Predictive modeling	Inference attacks

5.2 Implementation Effectiveness

Table 2 demonstrates privacy measure improvements following implementation.

Privacy Measure	Before (%)	After (%)	Improvement (%)
Data Encryption	60	90	30
Access Control	55	85	30
Anonymization	50	80	30
Data Masking	45	75	30

Implementation resulted in 40% reduction in data breach occurrences and 30% increase in user satisfaction, indicating enhanced confidence in privacy controls.

5.3 Regulatory Compliance

Table 3 presents compliance improvements across regulatory frameworks.

Regulatory Requirement	Before (%)	After (%)	Improvement (%)
HIPAA	65	95	30
GDPR	70	90	20
National Standards	60	85	25

5.4 Stakeholder Feedback

User satisfaction improved significantly across all measured criteria, with increases ranging from 0.9 to 1.4 points on a 5-point scale. Healthcare practitioners reported enhanced secure data management capabilities, while patients expressed increased confidence in organizational privacy protection measures [10,11].

VI. DISCUSSION

The evaluation reveals the complex balance required between privacy protection and data utility in healthcare systems. Implementation of systematic privacy-preserving measures achieved substantial effectiveness improvements while maintaining data accessibility for clinical and research applications.

Cryptographic approaches including homomorphic encryption and secure multi-party computation provided strong privacy guarantees suitable for high-sensitivity applications but with computational limitations restricting real-time use. Differential privacy emerged as a balanced solution offering formal privacy guarantees with reasonable computational

DOI: 10.17148/IJIREEICE.2025.13819

efficiency. Federated learning demonstrated particular promise for collaborative healthcare applications, enabling knowledge sharing without centralized data aggregation, though requiring additional safeguards against inference attacks. The 30% improvement in regulatory compliance across HIPAA, GDPR, and national standards confirms that advanced PPTs effectively address current legal requirements while providing flexibility for evolving regulations. However, regulatory fragmentation across jurisdictions creates implementation challenges for international healthcare collaborations.

User satisfaction improvements exceeding 1 point on a 5-point scale validate practical acceptance of privacy-preserving techniques. Healthcare practitioners appreciated enhanced data management capabilities, while patients expressed increased confidence in organizational privacy protection. The findings emphasize that hybrid approaches combining multiple techniques provide optimal solutions, as no single method excels across all evaluation criteria.

VII. CONCLUSION

This analysis demonstrates that systematic implementation of privacy-preserving techniques significantly enhances healthcare data security, regulatory compliance and stakeholder confidence while preserving data utility. The evaluation revealed 30% improvements in key privacy measures, 40% reduction in data breach occurrences and substantial regulatory compliance gains across HIPAA (30%), GDPR (20%) and national standards (25%).The comparative analysis confirms that traditional anonymization methods provide insufficient protection against modern attacks, while advanced techniques including homomorphic encryption, differential privacy and federated learning offer stronger privacy guarantees with varying trade-offs. Hybrid approaches combining multiple techniques provide optimal solutions for real-world healthcare environments, as no single method excels across all evaluation criteria. User satisfaction improvements exceeding 1.3 points validate the practical effectiveness and acceptance of privacy-preserving techniques in healthcare settings. The results demonstrate that robust privacy protection enhances rather than hinders healthcare operations, supporting broader adoption of data-driven healthcare innovations.

Future research should focus on developing standardized evaluation frameworks, simplified deployment tools and best practices for hybrid implementations. Healthcare organizations should adopt privacy-by-design principles, while policymakers should update regulatory frameworks to accommodate advanced privacy-preserving technologies. Success requires interdisciplinary collaboration among computer scientists, healthcare professionals, ethicists, and policymakers to ensure technical solutions address both practical healthcare needs and fundamental privacy rights.

REFERENCES

[1]. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst. 2014;2(1):3.
[2]. Krumholz HM. Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system. Health Aff (Millwood). 2014;33(7):1163–70.
[3]. Kruse CS, Frederick B, Jacobson T, Monticone DK. Cybersecurity in healthcare: A systematic review of modern threats and trends. Technol Health Care. 2017;25(1):1–10.
[4]. McGraw D, Dempsey JX, Harris L, Goldman J. Privacy as an enabler, not an impediment: building trust into health information exchange. Health Aff. 2009;28(2):416–27.
[5]. Department of Health and Human Services (HHS). Anthem Inc. breach report. 2015.
[6]. Guddati V, Guddati AK. Ethical issues in patient data ownership. Interact J Med Res. 2021;10. https://doi.org/10.2196/22269
[7]. Sullivan M, Howard R. Qualitative methods in privacy research. J of Infor Priv Sec. 2020;16(1):1-17.
[8]. Rocher L, Hendrickx JM, de Montjoye YA. Estimating the success of re-identifications in incomplete datasets using generative models. Nat Commun. 2019;10(1):3069.
[9]. HIPAA Journal. What is HIPAA compliance? 2023.
[10]. Cao S, Zhang XS, Xu RX. Toward secure storage in cloud-based ehealth systems: A blockchain-assisted approach. IEEE Network. 2020;34:64–70.
[11]. Kuo M, et al. Adoption of Software-as-a-Service (SaaS) solutions in healthcare: A systematic review. J Healthc Infor Man. 2020;34(2):45-56.
[12]. Houliston B, Shah MA, Nguyen H, Mather P. Privacy preserving healthcare data management: A review of the current landscape and a path to a secure and scalable architecture. Fut Gen Comp Sys. 2020;108:952-77.

