

**Federated Learning for Privacy-Preserving Healthcare Systems A Literature
Review**

Student Preparation

Abdulrahman Mohammed Aljamaan

Under the supervision of

د. محمد عنبر

2026

Table of Contents

Abstract	3
Introduction	4
Background	5
Evolution of Artificial Intelligence in Healthcare	6
Healthcare Data and Privacy Challenges	6
Federated Learning	6
Applications for Federated Learning in Healthcare	7
Problem Statement	7
Research Objectives	9
Research Scope	10
Literature Review	11
Research Methodology	16
Research Design	16
Data Sources	16
Search Strategy	17
Study Selection Criteria	18
Data Extraction	19
Data Analysis	19
Ethical Considerations	20
Limitations	20
Expected Outcomes	20
Gantt Chart	22
References	23

Abstract

Advanced medical diagnosis, disease prediction, individualized treatment, and clinical decision support are all made possible by artificial intelligence (AI), which has emerged as a game-changing technology in the healthcare industry. Large-scale patient data is crucial for these applications' machine learning model training. However, there are serious issues with patient privacy, data security, and regulatory compliance when sensitive healthcare data is collected and stored centrally. Due to ethical obligations and regulatory constraints, healthcare companies are frequently unwilling to disclose patient data, which restricts the availability of different datasets for creating correct AI models.

A potential distributed machine learning methodology called federated learning (FL) makes it possible for several healthcare facilities to work together to jointly train a global model without sharing raw patient data. Only locally trained model parameters are exchanged with a central server rather than sensitive datasets, thus lowering privacy hazards without sacrificing model performance. Federated Learning still has a number of drawbacks despite its benefits, such as communication cost, statistical variability among medical datasets, susceptibility to hostile attacks, and the problem of guaranteeing scalability and fairness among participating institutions. This study suggests a thorough evaluation of the literature on federated learning for healthcare systems that protect patient privacy. The study intends to compare privacy-preserving methods integrated with FL, evaluate recent research advancements, determine the advantages and disadvantages of current Federated Learning methodologies, and investigate the main issues influencing the implementation of FL in real-world healthcare settings. Additionally, the analysis highlights existing research gaps and explores potential future paths to enhance Federated Learning's security, effectiveness, and practical implementation in healthcare settings. The goal of this research is to give scholars, medical professionals, and legislators a thorough understanding of Federated Learning technologies, privacy-preserving strategies, and new developments that facilitate safe and cooperative AI in contemporary healthcare systems.

Keywords: Federated Learning, Healthcare Systems, Privacy Preservation, Artificial Intelligence, Distributed Machine Learning, Medical Data Security.

Introduction

Over the past ten years, artificial intelligence (AI) has transformed several industries, with healthcare emerging as one of the most important fields to gain from AI-driven breakthroughs. Disease diagnosis, medical image analysis, individualized therapy recommendations, drug development, patient monitoring, and clinical decision support systems have all significantly improved thanks to the application of machine learning and deep learning techniques in the healthcare industry. These developments have improved patient outcomes, decreased medical errors, and increased the effectiveness and precision of healthcare services.

The availability of sizable, varied, and high-quality datasets is crucial to the success of contemporary AI systems. Medical histories, diagnostic reports, test findings, medical photographs, and genetic information are among the extremely sensitive patient data found in these datasets. To train machine learning models, patient data from several hospitals and healthcare facilities must be gathered centrally. This centralized method raises serious problems about data privacy, security, confidentiality, and regulatory compliance even if it frequently increases model performance (Fahim, 2025).

One of the most sensitive types of personal information is healthcare data. Patients and healthcare providers may face serious ethical, legal, and financial repercussions from unauthorized access, data breaches, or abuse of medical records. Furthermore, the collection, storage, processing, and sharing of healthcare data are subject to stringent restrictions like the General Data Protection Regulation (GDPR) in Europe and the Health Insurance Portability and Accountability Act (HIPAA) in the United States. These rules frequently restrict the exchange of data between institutions, which makes it difficult to create reliable AI models that need big datasets. Federated Learning (FL), a cutting-edge distributed machine learning paradigm that facilitates cooperative model training without needing healthcare organizations to share raw patient data, has evolved to address these issues. Each participating institution trains the model locally using its own data, sharing only the model parameters or gradients rather than sending sensitive datasets to a centralized server. The original patient data is safely stored within each institution while a central server compiles these updates to create an enhanced global model.

Federated learning adoption has various benefits for healthcare applications. It facilitates cooperation between geographically dispersed hospitals and medical facilities, improves patient

privacy, lowers the dangers connected with centralized data storage, and helps regulatory compliance. Federated learning has therefore garnered significant interest in several healthcare applications, such as disease prediction, wearable health monitoring, electronic health record analysis, medical image classification, and cooperative clinical research (Ram, 2024).

Federated Learning has many drawbacks despite these encouraging benefits. Communication overhead between participating institutions, heterogeneity of medical datasets (non-IID data), disparities in computational resources, susceptibility to adversarial attacks, privacy leakage through model updates, and preserving equity across diverse healthcare organizations are just a few of the technical and operational obstacles that practical implementation must overcome. To facilitate the widespread implementation of Federated Learning in actual healthcare settings, addressing these issues is still a topic of current research.

This study offers a thorough analysis of the literature on Federated Learning for healthcare systems that protect patient privacy. Examining recent developments, contrasting current approaches, analyzing privacy-preserving strategies, identifying present constraints, and highlighting prospects for future study are all part of its objectives. This research aims to provide thorough knowledge of how Federated Learning might enable safe, cooperative, and reliable artificial intelligence in contemporary healthcare systems by combining insights from current scientific investigations (Ali M. S., 2024).

Background

Artificial intelligence's quick development and the healthcare industry's growing digital transformation have opened previously unheard-of possibilities for enhancing medical services through intelligent data analysis. However, there are now serious worries about patient privacy and data security due to the increased reliance on large-scale healthcare information. Examining the development of AI in healthcare, the features of healthcare data, current privacy issues, and the rise of distributed learning approaches as a privacy-preserving solution are all necessary to comprehend the history of federated learning (Bajwa, 2021).

Evolution of Artificial Intelligence in Healthcare

From rule-based expert systems to sophisticated machine learning and deep learning models that can resolve challenging healthcare issues, artificial intelligence has improved. Contemporary artificial intelligence (AI) systems are capable of autonomously analyzing medical images, forecasting the course of diseases, helping doctors diagnose patients, suggesting treatment options, and enhancing hospital resource management. The implementation of AI technologies in healthcare companies has accelerated due to the growing availability of wearable technology, genomic data, medical imaging datasets, and electronic health records (EHRs).

AI applications have spread to many medical areas, such as radiology, pathology, cardiology, oncology, ophthalmology, and intensive care medicine, as computing powers continue to advance. In certain diagnostic tasks, these applications have shown performance on par with, and occasionally better than, human experts (kumar, 2026).

Healthcare Data and Privacy Challenges

Because it includes financial, genetic, clinical, and personal information that is directly related to specific patients, healthcare data is extremely sensitive. Although accessibility and clinical decision-making have been enhanced by the extensive digitization of healthcare records, the potential of hacks, unauthorized data sharing, and privacy violations has also increased. Prior to model training, hospitals must move patient records to a central repository under traditional centralized machine learning. This strategy raises a number of issues, such as worries about data ownership, legal limitations, heightened cybersecurity risks, and a decline in healthcare organizations' desire to work together. As a result, many important healthcare datasets are still sold within certain institutions, which hinders the creation of precise and broadly applicable AI models.

Federated Learning

A networked machine learning system called Federated Learning was developed to facilitate cooperative model training while protecting data privacy. Each participating healthcare facility uses its own data to train the machine learning model locally rather than transferring patient data to a centralized server. A central aggregation server receives just the learnt model parameters or

gradients, combining updates from several participants to create an enhanced global model. While enabling institutions to gain from collaborative learning, this decentralized training approach drastically lowers the exposure of private medical information. Federated Learning has therefore emerged as one of the most promising privacy-preserving medical AI solutions.

Applications for Federated Learning in Healthcare

Numerous healthcare applications, such as medical image analysis, cancer detection, diabetic retinopathy diagnosis, COVID-19 detection, brain tumor segmentation, predictive analytics using electronic health records, wearable healthcare devices, remote patient monitoring, and personalized medicine, have effectively investigated federated learning.

By utilizing information from several healthcare institutions without jeopardizing patient privacy, these applications show how collaborative learning can enhance model generalization. As a result, federated learning is becoming more widely acknowledged as a crucial technology for delivering safe and reliable AI-powered healthcare systems (Wassan, 2025).

Problem Statement

By enabling intelligent diagnostic tools, predictive analytics, medical picture interpretation, individualized therapy recommendations, and clinical decision support, the quick development of artificial intelligence (AI) has profoundly changed the healthcare industry. Large-scale healthcare datasets are crucial to these AI systems' high predicted accuracy and dependability. Traditionally, gathering patient data from many healthcare facilities into centralized repositories has been necessary for the development of such models. However, there are significant issues with patient privacy, data confidentiality, cybersecurity risks, and adherence to healthcare standards with this centralized method.

Because of privacy laws like the Health Insurance Portability and Accountability Act (HIPAA), the General Data Protection Regulation (GDPR), and institutional data governance principles, healthcare institutions are sometimes reluctant or legally prohibited from sharing sensitive patient information. As a result, important medical data is still dispersed among hospitals and healthcare

providers, which restricts the availability of large datasets needed to train reliable AI models. This data separation limits collaborative healthcare innovation, slows medical research, and decreases model generalization.

A promising distributed learning paradigm that permits cooperative model training without sending raw patient data outside of local institutions is federated learning (FL). Federated Learning greatly enhances privacy protection and facilitates regulatory compliance by enabling healthcare firms to train models locally while sharing only model parameters. Federated Learning has a lot of potential, but there are several operational and technical issues that prevent it from being widely used in the healthcare industry.

Communication overhead during collaborative training, statistical heterogeneity resulting from non-identically distributed (non-IID) medical datasets, disparities in computational resources between participating institutions, model convergence challenges, susceptibility to adversarial attacks, potential information leakage through model updates, and preserving fairness across heterogeneous healthcare environments are some of these challenges. Federated learning systems' security and performance are both impacted by these problems, which are still being studied. There is still a dearth of thorough research contrasting privacy-preserving strategies, secure aggregation protocols, differential privacy methods, and blockchain-integrated Federated Learning frameworks in healthcare applications. Additionally, current literature frequently concentrates on specific technical features without offering a comprehensive knowledge of their advantages, disadvantages, real-world implementation difficulties, and prospects for future research.

Therefore, by analyzing recent research contributions, assessing current methodologies, identifying current research gaps, and highlighting future directions that can facilitate the safe, effective, and reliable deployment of Federated Learning in contemporary healthcare environments, this study seeks to conduct a thorough literature review of Federated Learning for privacy-preserving healthcare systems (Faiyazuddin, 2025).

Research Objectives

The primary objective of this research is to conduct a comprehensive literature review on Federated Learning for privacy-preserving healthcare systems and evaluate its effectiveness in enabling secure collaborative artificial intelligence while maintaining patient privacy.

To achieve this objective, the research aims to accomplish the following objectives:

1. To review the evolution of Federated Learning in healthcare.

- Examine the fundamental concepts and architecture of Federated Learning.
- Analyze the motivation behind adopting Federated Learning in healthcare environments.
- Discuss the advantages of decentralized learning over traditional centralized machine learning.

2. To analyze privacy and security challenges in healthcare AI.

- Identify major privacy concerns associated with centralized healthcare data.
- Examine cybersecurity risks affecting healthcare machine learning systems.
- Discuss legal and ethical considerations related to patient data sharing.

3. To evaluate existing Federated Learning approaches for healthcare.

- Compare different Federated Learning architectures and aggregation algorithms.
- Evaluate privacy-preserving techniques integrated with Federated Learning.
- Analyze recent applications across various medical domains.

4. To identify the strengths and limitations of current Federated Learning techniques.

- Investigate communication efficiency challenges.
- Analyze data heterogeneity (Non-IID) problems.
- Examine scalability, fairness, robustness, and model performance.
- Evaluate vulnerabilities to adversarial attacks and privacy leakage.

5. To identify research gaps and future research directions.

- Highlight limitations in current literature.
- Identify emerging technologies supporting Federated Learning.
- Suggest future research opportunities for improving privacy-preserving healthcare AI systems.

Research Scope

The overview and analysis of current research on Federated Learning for privacy-preserving healthcare systems is the main goal of this study. The work focuses on comprehending how Federated Learning permits collaborative AI while maintaining patient privacy and adhering to healthcare privacy laws.

The scope of this research includes:

- going over Federated Learning's basic principles, architecture, and process.
- investigating the use of federated learning in healthcare applications, such as clinical decision support systems, wearable medical technology, electronic health records, medical imaging, and disease diagnostics.
- examining privacy-preserving methods such Blockchain-based Federated Learning, Differential Privacy, Secure Aggregation, and Homomorphic Encryption.
- contrasting contemporary literature's suggested security measures, communication tactics, aggregation algorithms, and Federated Learning frameworks.
- identifying the main technological obstacles to Federated Learning deployment, such as adversarial assaults, scalability, non-IID data distribution, communication overhead, and model convergence.
- assessing current research trends, pointing out research gaps, and talking about potential future paths to enhance safe collaborative learning in healthcare.

Federated Learning algorithm implementation and experimental evaluation are not included in this study, which is restricted to a survey of the literature. Rather, it provides a thorough grasp of the

state of the art in privacy-preserving Federated Learning for healthcare systems by synthesizing findings from published scientific publications.

Literature Review

A promising technique for training machine learning models while maintaining patient privacy is federated learning (FL). The application of privacy-preserving strategies in FL within decentralized healthcare systems is examined in this systematic review. Based on regulatory compliance, scalability, computational cost, complexity, and mathematical underpinnings, it evaluates current techniques like Differential Privacy (DP), Trusted Execution Environment (TEE), Zero Knowledge Proofs (ZKP), Homomorphic Encryption (HE), Watermarking, Blockchain, and Secure Multi-Party Computation (SMPC). There has been discussion of the main difficulties with decentralized healthcare, including diverse data, privacy hazards, security threats, and compliance issues. The assessment also emphasizes how crucial it is to abide by national data protection rules as well as international requirements like GDPR and HIPAA. Additionally, it addresses unresolved issues and recommends future lines of inquiry to get around present constraints, such as adversarial attacks, computational efficiency, and the development of standardizing policy frameworks. For the upcoming generation of healthcare applications, this paper offers a distinctive viewpoint on scalable, secure, and ethical privacy-preserving FL models (Kumar, 2025).

Healthcare organizations can cooperatively train machine learning models without disclosing private patient information by using federated learning (FL). By combining privacy protection, algorithmic fairness, governance, and equitable access into a single analytical framework, this systematic review seeks to fully integrate the ethical aspects of FL in healthcare. With an emphasis on ethical concerns including algorithmic fairness, privacy protection, governance, and equitable access, the use of FL in healthcare between January 2020 and December 2024 is analyzed. Methods: Six databases (PubMed, IEEE Xplore, Web of Science, Scopus, ACM Digital Library, and arXiv) were searched in accordance with PRISMA criteria. CRD420251274110 is the PROSPERO registration number. Studies that specifically addressed ethical issues and detailed FL deployments in healthcare contexts were chosen. FL architectures, privacy-preserving techniques

like encryption, safe multiparty computation, and differential privacy, as well as fairness measures, governance models, and clinical application fields, were among the important data that were retrieved. Results: 38 of the 3047 recordings satisfied the requirements for inclusion. Electronic health records and medical imaging, particularly in radiology and oncology, were the most widely used applications. Four major ethical themes were identified through thematic analysis: algorithmic fairness, which considers client and attribute differences; privacy protection through formal guarantees and cryptographic techniques; governance models, which prioritize accountability, transparency, and stakeholder engagement; and equitable distribution of computing resources for institutions with limited resources. Only a few research documented real-world clinical deployment, and there was significant heterogeneity in the evaluation of privacy and fairness trade-offs. Conclusions: FL offers a great deal of potential to advance ethical AI in healthcare, but progress will necessitate the creation of shared fairness standards, practical governance strategies, and mechanisms to ensure equitable benefit sharing. Standardized fairness measures, multi-stakeholder governance frameworks, and real-world clinical validation beyond proof-of-concept implementations should be the focus of future research (Mir, 2026).

An architecture for privacy-preserving federated learning that can categorize chest X-ray photos of tuberculosis (TB) in decentralized healthcare settings. By enabling the collaborative training of ML models across numerous healthcare institutions without requiring direct access to private patient data, the proposed method protects patient privacy and guarantees regulatory compliance. Initial inspection, picture resizing to 224x224 pixels, normalization, and class balancing utilizing Synthetic Minority Oversampling Technique (SMOTE) to handle non-IID distributions among federate clients are all part of a structured data pre-processing pipeline. Each node in the federated arrangement stores a subset of the dataset and only communicates model modifications to the central server, simulating realistic clinical settings. The ResNet deep learning model is used as the main classifier, and four important assessment metrics—accuracy, precision, recall, and F1-score—are used to compare its performance with Dense Net and Squeeze Net. The proposed ResNet model outperforms the comparable models in terms of accuracy (96.7%), precision (96.8%), recall (98.0%), and F1-score (97.4%). Accuracy rates for Squeeze Net and Dense Net were 94.18% and 94%, respectively. The scalability and real-time applicability of this framework are enhanced by its interaction with cloud-based platforms. Using federated learning and big data analytics, it provides a high-performance, scalable, and secure method for diagnosing tuberculosis

(TB) in healthcare settings. The outcome confirms its promise as a basis for further extensive uses in medical AI systems that respect privacy (Orthi, 2025).

By removing the need for data sharing, federated learning (FL) offers a distributed machine learning system that allows users to train using local data to produce a shared model. FL enables local training of electronic health records (EHRs) and Medical Internet of Things (MIoT) devices in healthcare systems without transmitting patient data to a central server. In addition to simplifying other healthcare procedures, this enables healthcare judgments and diagnoses based on datasets from all participants. This technique enables collaborative training without requiring the sharing of local data with the central server, thereby protecting user privacy. However, there are privacy issues with FL because the client and server communicate model changes, which can be exploited to re-generate the client's data and violate privacy regulations for applications in industries like healthcare. To analyze the current privacy and security enhancement techniques suggested for FL in healthcare systems, we have reviewed the literature in this research. Differential privacy, homomorphic encryption, blockchain, hierarchical approaches, peer-to-peer sharing, intelligence on the edge device, and mixed, hybrid, and miscellaneous approaches are the seven methods that have been found to be the focus of the field's research. Each technique's advantages, disadvantages, and trade-offs were examined, and the potential future of these seven privacy-enhancing methods for healthcare FL systems was determined (Gu, 2023).

By utilizing medical technologies, recent developments in electronic equipment and communication infrastructure have transformed the conventional healthcare system into a smart healthcare system. However, mobile and wearable IoMT devices present privacy concerns about the information shared between hospitals and end users because of the centralized training strategy of artificial intelligence (AI). Adversaries may be able to access the extremely private data that IoMT devices transmit. In this sense, federated learning (FL), a distributive AI paradigm, has created new potential for IoMT privacy preservation without gaining access to members' private information. Additionally, since only gradients are transmitted during training, FL offers end users privacy. In this study, we discuss privacy concerns in IoMT for these particular features of FL. Next, we discuss how FL protects privacy in IoMT networks and describe several sophisticated FL architectures that use generative adversarial networks (GANs), digital twins, and deep reinforcement learning (DRL) to identify privacy risks. Additionally, we offer several real-world

applications for FL in IoMT. Finally, we wrap up this study by talking about challenges and open research questions related to the use of FL in future smart healthcare systems (Ali M.).

The need for frameworks for data exchange while maintaining patient privacy and safety has been sparked by the healthcare industry's rapid digital transformation in the age of artificial intelligence and the Internet of Medical Things. Centralized learning approaches compromise responsiveness and responsible ownership of huge medical data by putting potentially sensitive patient data at risk of leakage, legal violations, and cyberattacks. Federated learning is a new paradigm that lets patients from different healthcare organizations train machine learning models without having to share their own data. A thorough evaluation of the literature on privacy-preserving federated learning framework approaches in healthcare applications is suggested by this study. Web of Science, Scopus, IEEE Xplore, ScienceDirect, PubMed, and the ACM Digital Library were searched using predetermined query strings, clear inclusion/exclusion criteria, and quality appraisal techniques in accordance with PRISMA recommendations. This systematic review, which looked at cryptographic, architectural, and algorithmic techniques like differential privacy, homomorphic encryption, and Secure Multi-Party Computation, as well as blockchain integrations to improve trust and confidence in distributed healthcare systems, included 80 peer-reviewed studies published between January 2015 and December 2025. The results show a slow transition to hybrid privacy-preserving federated learning architectures that incorporate several security measures to increase resilience, confidentiality, and trust. Even though there has been a lot of development, the difficulties with communication efficiency, non-IID data dissemination, adversarial assaults, and regulatory restrictions have a big impact on the real-world implementation of such systems. To achieve the best possible balance between privacy, utility, and system performance for next-generation health intelligence, this study identifies future research areas for scalable, explainable, and interoperable federated architectures (Alankamony, 2026).

Institutional reluctance to share patient data, scattered data sources, and privacy laws frequently make it difficult to integrate artificial intelligence (AI) into medical imaging. A promising method that allows for collaborative model training without the necessity for data centralization is federated learning (FL). The function of FL in medical imaging is critically examined in this paper, with an emphasis on privacy-preserving methods. We examine advancements in FL designs, privacy-enhancing methods, and medical imaging-specific problems such data heterogeneity,

modality-specific limitations, communication overhead optimization tactics, and modality drift-related fairness concerns. The review also discusses the effects of international health regulations, such as India's Digital Personal Data Protection (DPDP) Act, the General Data Protection Regulation (GDPR), and the Health Insurance Portability and Accountability Act (HIPAA), and offers insights into various privacy-preserving strategies. The overall goal of this review is to direct the creation of practical, safe, and privacy-conscious FL systems that are suited for actual medical imaging applications (Pechetti, 2026).

An age of significant change is being ushered in by the field of healthcare data collaboration, which highlights a remarkable potential to improve patient care and accelerate medical research. However, protecting the security and privacy of sensitive data is a huge effort that presents many challenges. Using privacy-preserving federated learning models, this paper offers a creative solution to these problems, thereby paving the way for a new direction in this complex area of study. Our suggested approach protects the privacy of individual patient data while allowing healthcare organizations to jointly train machine learning models on decentralized data. The suggested mechanism integrates state-of-the-art privacy-preserving techniques, such as differential privacy and secure multi-party computation, to enforce the security of sensitive data during the model aggregation phase. We do a variety of thorough simulations and assessments with a focused focus on accuracy, computational efficiency, and privacy preservation in order to validate the effectiveness of the suggested solution. The findings collected confirm that our methodology outperforms competing alternatives in terms of higher utility and strong privacy guarantees. The suggested method is a convincing example of its usefulness and efficacy since it captures the viability of safe and privacy-preserving cooperation on medical data. Through our work, we confirm the promise of a new era in collaborative healthcare informatics by highlighting the possibility of utilizing collective intelligence in healthcare while upholding the highest standards of privacy protection (Abaoud, 2023).

By using data to create models that may guide clinical operations, artificial intelligence (AI) has the potential to enhance healthcare. However, building reliable, broadly applicable models requires access to vast amounts of heterogeneous data. Because of privacy, security, and legal issues, data exchange between institutions is not always possible. Although it raises different security and privacy issues, federated learning (FL) enables multi-institutional training of AI models,

eliminating the need for data exchange. In particular, information concerning institutional data may be revealed by insights shared during FL. Furthermore, when there is a lack of confidence between the entities carrying out the computation, FL may cause problems. Clarifying the possible hazards is crucial given the increasing use of FL in healthcare. In this paper, we thereby provide a summary of privacy-preserving FL literature with a focus on health care. We highlight risks and examine ways to mitigate them. We hope that our analysis will serve as a resource for healthcare researchers in Florida about security and privacy (Pati, 2024).

Research Methodology

Research Design

Federated Learning for privacy-preserving healthcare systems is thoroughly examined in this study using a Systematic Literature Review (SLR) technique. Because it makes it possible to identify, assess, and synthesize current scientific information in an organized, transparent, and repeatable way, a systematic literature review was chosen. In contrast to traditional narrative reviews, the SLR technique minimizes bias and increases the reliability of the results by following predetermined procedures for identifying, choosing, and evaluating pertinent papers.

Recent advancements in Federated Learning, privacy-preserving strategies, healthcare applications, security issues, and potential future research areas are the main topics of this paper. By using this method, the study seeks to uncover existing knowledge gaps that need more research while offering a current picture of current research trends (Mohammadi, 2024).

Data Sources

Relevant studies were collected from internationally recognized digital scientific databases to ensure the inclusion of high-quality peer-reviewed publications. The primary databases considered in this research include:

- IEEE Xplore Digital Library
- ScienceDirect

- SpringerLink
- ACM Digital Library
- PubMed
- Scopus
- Web of Science
- Google Scholar

These databases provide extensive coverage of computer science, artificial intelligence, cybersecurity, and healthcare research, making them appropriate sources for identifying publications related to Federated Learning and privacy-preserving healthcare systems.

Search Strategy

A structured keyword-based search strategy was employed to identify relevant studies. Different combinations of keywords were used to maximize the retrieval of related publications.

The primary search keywords included:

- Federated Learning
- Privacy-Preserving Machine Learning
- Healthcare Artificial Intelligence
- Medical Data Privacy
- Electronic Health Records
- Medical Imaging
- Distributed Machine Learning
- Differential Privacy
- Secure Multi-Party Computation
- Homomorphic Encryption

- Blockchain in Healthcare
- AI Security

Boolean operators such as **AND**, **OR**, and **NOT** were used to refine search results and improve retrieval accuracy.

Study Selection Criteria

To ensure the quality and relevance of the reviewed literature, predefined inclusion and exclusion criteria were established.

Inclusion Criteria

The selected studies satisfied one or more of the following conditions:

- Published in peer-reviewed journals or reputable international conferences.
- Written in English.
- Published between 2017 and 2026.
- Focused on Federated Learning in healthcare.
- Addressed privacy-preserving techniques or healthcare security.
- Presented experimental evaluation, systematic review, or comprehensive framework.

Exclusion Criteria

The following studies were excluded:

- Non-English publications.
- Duplicate articles.
- Studies unrelated to healthcare.
- Studies focusing solely on traditional centralized machine learning.
- Editorials, short abstracts, and non-peer-reviewed publications.

Data Extraction

For each selected study, the following information was extracted and systematically recorded:

- Authors
- Publication year
- Research objective
- Healthcare application
- Federated Learning approach
- Privacy-preserving technique
- Dataset
- Evaluation metrics
- Main findings
- Strengths
- Limitations

This information facilitated a structured comparison among different studies and enabled the identification of common research trends.

Data Analysis

The collected studies were analyzed using a qualitative comparative analysis approach. The analysis focused on identifying recurring themes, similarities, differences, strengths, limitations, and emerging research trends across the selected publications.

The reviewed studies were categorized into several major themes, including:

- Federated Learning architecture.
- Privacy-preserving techniques.

- Security mechanisms.
- Healthcare applications.
- Communication efficiency.
- Non-IID data handling.
- Regulatory compliance.
- Future research directions.

Comparative analysis enabled a comprehensive understanding of the current state of Federated Learning in healthcare while identifying research gaps that require further investigation.

Ethical Considerations

Clinical investigations, patient data, or human subjects were not used in this study. Thus, ethical approval was not necessary. To maintain academic integrity and prevent plagiarism, all of the data used in this work came from publically accessible scientific papers and were properly cited.

Limitations

Several limitations should be noted, even though this systematic literature review offers a thorough overview of Federated Learning in healthcare. First, only English-language studies were considered, which would have left out pertinent material that was available in other languages. Second, the review may have missed important technical papers or industrial implementations because it mainly concentrated on peer-reviewed publications that were indexed in significant scientific databases. Lastly, because Federated Learning is always changing, new research may be released after this evaluation is finished.

Expected Outcomes

It is anticipated that this study will offer a thorough and current understanding of Federated Learning (FL) as a privacy-preserving machine learning paradigm for contemporary healthcare

systems. The study intends to provide a clear overview of the development of Federated Learning, its underlying architectures, and its expanding role in enabling secure collaboration among healthcare institutions without requiring the exchange of sensitive patient data by methodically reviewing and analyzing recent scientific literature. It is anticipated that the study will assess the efficacy of several privacy-preserving methods combined with Federated Learning, such as Differential Privacy, Homomorphic Encryption, Secure Multi-Party Computation, Blockchain, Trusted Execution Environments, and other cutting-edge security mechanisms, while analyzing their advantages, disadvantages, and suitability for use in actual healthcare settings. It is also expected that the study will identify the main operational and technical obstacles to Federated Learning deployment, including communication overhead, non-IID data distribution, computational complexity, scalability, interoperability, fairness, vulnerability to adversarial attacks, and privacy leakage. The study is anticipated to identify opportunities for future research aimed at enhancing the effectiveness, robustness, explainability, and security of Federated Learning frameworks, as well as to reveal current research trends and highlight important gaps in literature through a comparative analysis of existing studies. Ultimately, by acting as a trustworthy resource that promotes the creation and uptake of safe, scalable, moral, and privacy-preserving artificial intelligence solutions capable of enhancing healthcare services while guaranteeing compliance with data protection laws and upholding patient trust, this literature review is anticipated to offer insightful information to researchers, healthcare professionals, developers, and legislators (Ajax, 2025).

Gantt Chart

Tasks		Sep 2025	Oct 2025	Nov 2025	Dec 2025	Jan 2026	Feb 2026
	Topic Selection						
	Literature Search & Review						
	Collect Relevant Papers						
	Analyze & Extract Data						
	Write Literature Review						
	Research Methodology						
	Expected Outcomes						
	Compile Proposal						
	Editing & Proofreading						
	Final Submission						

References

- Abaoud, M. (2023). Advancing Federated Learning Through Novel Mechanism for Privacy Preservation in Healthcare Applications. Retrieved from <https://ieeexplore.ieee.org/abstract/document/10201859#citations>
- Ajax, R. (2025). Federated Learning for Privacy-Preserving Healthcare Analytics: Securing Patient Data through Distributed Machine Learning. Retrieved from https://www.researchgate.net/publication/393790854_Federated_Learning_for_Privacy-Preserving_Healthcare_Analytics_Securing_Patient_Data_through_Distributed_Machine_Learning
- Alankamony, A. (2026). Systematic Review of Privacy Preservation in Federated Learning for Secured Healthcare Applications. Retrieved from <https://www.mdpi.com/2078-2489/17/7/647>
- Ali, M. (n.d.). Federated Learning for Privacy Preservation in Smart Healthcare Systems: A Comprehensive Survey. 2023. Retrieved from <https://ieeexplore.ieee.org/document/9794622>
- Ali, M. S. (2024). Federated Learning in Healthcare: Model Misconducts, Security, Challenges, Applications, and Future Research Directions -- A Systematic Review. Retrieved from https://www.researchgate.net/publication/380820175_Federated_Learning_in_Healthcare_Model_Misconducts_Security_Challenges_Applications_and_Future_Research_Directions_--_A_Systematic_Review
- Bajwa, J. (2021). Artificial intelligence in healthcare: transforming the practice of medicine. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC8285156/>
- Fahim, Y. A. (2025). Artificial intelligence in healthcare and medicine: clinical applications, therapeutic advances, and future perspectives. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC12455834/>
- Faiyazuddin, M. (2025). The Impact of Artificial Intelligence on Healthcare: A Comprehensive Review of Advancements in Diagnostics, Treatment, and Operational Efficiency. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC11702416/>

- Gu, X. (2023). A Review of Privacy Enhancement Methods for Federated Learning in Healthcare Systems. Retrieved from <https://www.mdpi.com/1660-4601/20/15/6539>
- Kumar, K. S. (2025). Systematic review of privacy-preserving Federated Learning in decentralized healthcare systems. Retrieved from <https://www.sciencedirect.com/science/article/pii/S2773186325002257>
- kumar, N. (2026). Artificial Intelligence: Revolutionizing the Future of Medicine: Intelligent Systems, Deep Learning, and the Next Generation of Clinical Care. Retrieved from https://www.researchgate.net/publication/404316033_Artificial_Intelligence_Revolutionizing_the_Future_of_Medicine_Intelligent_Systems_Deep_Learning_and_the_Next_Generation_of_Clinical_Care
- Mir, B. A. (2026). Federated Learning in Healthcare Ethics: A Systematic Review of Privacy-Preserving and Equitable Medical AI. Retrieved from <https://www.mdpi.com/2227-9032/14/3/306>
- Mohammadi, S. (2024). Balancing Privacy and Performance in Federated Learning: a Systematic Literature Review on Methods and Metrics. Retrieved from https://www.researchgate.net/publication/380672479_Balancing_Privacy_and_Performance_in_Federated_Learning_a_Systematic_Literature_Review_on_Methods_and_Metrics
- Orthi, S. M. (2025). Federated Learning with Privacy-Preserving Big Data Analytics for Distributed Healthcare Systems. Retrieved from <https://alkindipublisher.com/index.php/jcsts/article/view/10509>
- Pati, S. (2024). Privacy preservation for federated learning in health care. Retrieved from [https://www.cell.com/patterns/fulltext/S2666-3899\(24\)00082-5](https://www.cell.com/patterns/fulltext/S2666-3899(24)00082-5)
- Pechetti, U. S. (2026). Privacy-Preserving Federated Learning in Medical Imaging: A Systematic Review. Retrieved from <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.70109>
- Ram, A. (2024). Artificial Intelligence in Healthcare: Revolutionizing Disease Diagnosis and Treatment Planning. Retrieved from

https://www.researchgate.net/publication/382273650_Artificial_Intelligence_in_Healthcare_Revolutionizing_Disease_Diagnosis_and_Treatment_Planning

Wassan, S. (2025). Federated learning and differential privacy: Machine learning and deep learning for biomedical image data classification. Retrieved from https://www.researchgate.net/publication/395416405_Federated_learning_and_differential_privacy_Machine_learning_and_deep_learning_for_biomedical_image_data_classification