International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 6s, 2025
https://www.theaspd.com/ijes.php

AI-POWERED HEALTH CARE: BALANCING DATA UTILITY AND PATIENT PRIVACY

Dr. Prachi Tripathi, Assistant Professor, Department of ECE,
Noida Institute of Engineering and Technology, 19, knowledge Park-II,
Greater Noida, Uttar Pradesh - 201 306 ,dhoolika77@gmail.com

Dr. Reshma M, Assistant Professor, Department of Electronics and communication Engineering,
University BDT college of Engineering, Karnataka – 577004,reshma.m03@gmail.com

Yalamandeswara Rao Gumma, Assistant Professor & Science and Humanities,
Vignan Pharmacy College, Vadlamudi

Tanneru Venkata Lavanya, Assistant Professor, Department of CSE, MALLA REDDY Vishwavidyapeeth Deemed to be University(Technical Campus), Secunderabad, Telangana, vlavanyat@gmail.com

Dr. Anita Pradhan, Assistant Professor, Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur,
Andhra Pradesh- 522502,anita.pradhan15@gmail.com

Abstract
Artificial intelligence (AI) has changed the landscape of healthcare systems, including medical diagnostic, treatment planning, and patient monitoring. Nonetheless, such a fast development provokes serious questions about patient privacy, especially in the age of big data and electronic health records. In this paper, the authors explore the twofold problem of making the most AI models useful with the data and ensuring the privacy of patients. It summarizes recent advances in privacy-preserving methods including differential privacy, federated learning and homomorphic encryption. A comparative analysis and a prototype implementation performed in the course of the study show that privacy-enhancing technologies can reduce the risks but there exists a trade-off between the model accuracy and the complexity of the resulting system. The study provides a conclusion with the suggestion of a balanced framework, which maximizes the utility of the data and privacy guarantees of AI-driven healthcare applications.

Keywords– *AI in Healthcare, Patient Privacy, Data Utility, Differential Privacy, Federated Learning, Privacy-Preserving Machine Learning, Medical AI, Data Ethics.*

INTRODUCTION

Over the past few years, Artificial Intelligence (AI) started to transform the contemporary healthcare sector, providing it with smarter, faster, and more accurate interventions. Whether it is the AI-powered diagnostic solution that interprets radiographic images at the expert level or the predictive models that can identify possible health hazards prior to their symptoms appearing, the application of machine learning to clinical practice is transforming the nature of healthcare delivery. The presence of vast amounts of healthcare data, such as electronic health records (EHRs), wearable device data, medical imaging, and genomics have driven most of these innovations. AI systems, particularly those based on deep learning and other data-hungry algorithms, need access to such data to enhance model training, personalization and general performance [9].But it poses a paradox as there is a greater reliance on patient data to develop AI models. Although more diverse data can improve the performance of AI, it leads to serious privacy implications. Health data comprises some of the most sensitive personally-identifying information, whose unauthorized access or misuse may have irreparable effects, including not only insurance discrimination and social stigma but also violations of confidentiality and loss of trust between patient and provider [10]. The long-established data anonymization techniques are becoming insufficient despite the implementation of various regulatory frameworks, such as HIPAA (Health Insurance Portability and Accountability Act) and GDPR (General Data Protection Regulation). Our ability to deal with such advanced re-identification attacks where anonymized data can be cross-referenced with other external data sets to de-anonymize personal information has rendered the need to

develop more rigorous privacy-preserving mechanisms to be immediate.Besides, data ownership and consent are emerging as key topics in ethical AI use in healthcare. Patients are becoming more eager to have control over the usage and sharing of their data, but the needs of such patients and the technicalities of developing a model often create a conflict that healthcare institutions and AI developers have to address [7]. More often than not, the lack of explicit regulations regarding the use of data during AI training has created ethical grey zones, which evoked the discussion of transparency, accountability, and informed consent.In these tensions, privacy-preserving technologies, including measures like differential privacy, federated learning, and homomorphic encryption, have been proposed as potentially great solutions. Differential privacy keeps statistical noise in databases, such that no particular record can be reconstructed without substantially altering the big-picture trends in the data. Instead, federated learning decentralizes model training by storing patient data locally and being selective of the model updates, which are sent to a central server. Homomorphic encryption enables calculations on encrypted data without decryption, which keeps privacy when processing data [8]These methods are technically possible, but both of them present a set of specific problems. As an example, differential privacy can cause the model accuracy to decrease as a result of injecting noise [2]. Federated learning involves considerable communication infrastructure, and suffers problems of model drift and system heterogeneity. Homomorphic encryption is theoretically secure but in most cases, significant computing resources are required which is not feasible to implement in large scale and in real-time clinical environment. Hence, the problem of finding the optimal compromise between the usefulness of data and the degree of privacy protection cannot be considered solely a technological issue-it is also a multidisciplinary one, touching upon law, ethics, computer science, and the practice of healthcare.The given paper will explore this fine balance by assessing the available privacy-preserving methods within the framework of AI-driven healthcare systems [4]. The study is expected to contribute to the provision of a practical guide to turning these techniques on model performance, user confidence, and complexity of implementation in an attempt to provide a practical guide to achieving high data utility and good patient privacy. The results have particular significance to policymakers, AI designers, hospital information technology managers, and clinicians who have to balance the two competing demands of innovation and safety in the era of digital health.

Novelty and Contribution

The proposed research has a number of new contributions to the emerging area of privacy-aware AI in healthcare. As opposed to the earlier studies, which frequently consider privacy and utility separately, the proposed paper provides an end-to-end hybrid framework to integrate two state-of-the-art privacy-preserving techniques, namely, differential privacy and federated learning, into a joint system. In such a way, the research not only improves the protection of data but also preserves satisfactory values of model performance in various artificial intelligence (AI) models, including logistic regression, convolutional neural networks (CNNs), and gradient-boosted trees [5].The other main novelty is the empirical assessment of privacy-utility trade-offs with the help of the real-world inspired healthcare data. The analysis is not limited to theoretical discourse, but the researchers provide quantitative exercises in which model accuracy, training overhead, and the probability of data leakage are compared with different privacy setups. It is a data-driven study that gives practical information to organizations seeking to apply such technologies in practice.The study further develops a user trust measure, which is ascertained by obtaining survey responses of healthcare users as they engage with prototype interfaces [3]. This humanistic view of design is usually lacking in technical studies yet is important to comprehend adoption issues in the real world. The paper provides a multidimensional evaluation framework of AI implementation in healthcare through the inclusion of technical, ethical, and user-experience dimensions.Last but not least, the paper provides a futuristic perspective on scalability, regulatory, and ethical control of AI in data-critical contexts. It provides some guidance on how privacy-preserving techniques can be incorporated into the institutional processes, how technical solutions can be coordinated with regulatory requirements, and how to make patients trust AI-empowered systems. The contributions made in combination can help not only enhance academic investigation but also create a guide to future-ready and privacy-preserving healthcare AI applications [16-17].

RELATED WORKS

In 2025 E. Gkiolnta et.al. D. Roy et.al. and G. F. Fragulis et.al. [15] suggested the intersection between artificial intelligence and healthcare has triggered widespread research on how to streamline the clinical outcomes using data-

driven approaches. The key element of this progress is using patient data to educate advanced models that can assist in diagnosis, forecasting treatment success, and enabling personalized medicine. Nevertheless, the outstanding necessity of securing sensitive patient data has also driven a parallel rise in the studies on privacy-preserving methods and responsible data handling in AI-based systems.Research in utility of data in healthcare owes that machine learning algorithms, particularly deep learning models, show their best performance when trained on large, diverse and high quality datasets. Whether it is electronic health records (EHRs), radiology scans, or genomics, clinical data provides useful patterns on which predictive analytics can be informed. However, health data are usually fragmented across organizations, irregular in structure, and tightly controlled in access because of privacy policies. These are impediments, which restrict the complete use of AI potential. Therefore, balancing data availability and data confidentiality has become an important topic of investigation.Data anonymization and pseudonymization privacy-preserving methodologies have been traditionally applied to obscure personally identifiable information. These techniques do provide some level of protection, however, they are becoming regarded as insufficient against modern re-identification techniques. Even anonymized dataset can be linked with other external data sources using sophisticated algorithms, thus pointing to the weakness of traditional de-identification measures.Due to the weakness of the simple anonymization, more complex methods have been investigated. Differential privacy adds random noise to the outputs of data, such that the presence or absence of any single datum point does not meaningfully change the analysis as a whole. The method has received interest due to its provable guarantees of privacy that are mathematically sound, though it may lower the accuracy of the model when not applied with careful tuning. The utility-noise trade-off is of special concern in healthcare, where precision may directly affect the outcome of patients [11].Another major technique that has come forth is federated learning to maintain privacy. It allows training machine learning models on decentralized devices or servers that each contain local data samples, without raw data communication. The architecture is particularly applicable in the healthcare sector in which institutions are usually unwilling to share patient records because of legal and ethical issues. Federated learning enables joint model construction without moving data beyond institutional premises. Nonetheless, this approach has problems, including communication overhead, model synchronization, and susceptibility to poisoning attacks in which malicious parties control the model.In 2024 S. M. Williamson et.al. and V. Prybutok et.al., [1] introduced the homomorphic encryption is a cryptographic breakthrough which allows computations to be performed on encrypted data, producing encrypted output that can be decrypted in the future to produce the answer.. Theoretically, it provides the best security because the raw data is not revealed at any point in processing. Nonetheless, it is yet to be applied in healthcare AI broadly because computation with the existing encryption schemes is costly and delayed. That has prompted ongoing efforts to increase the performance of homomorphic algorithms to make them practical.Another privacy-enhancing technique that has received momentum is synthetic data generation. Synonymous data Synthetic data can be used to train AI models without revealing patient data by generating artificial datasets that statistically resemble real patient data. Although encouraging, fidelity and generalizability of synthetic data pose a problem. It is also possible that badly produced synthetic data may incorporate subtle biases or structure of the original data, accidentally compromising privacy or damaging model outcomes.More research has been done on the social, ethical, and legal ramifications of AI in healthcare. Among researchers there is a push towards transparency, accountability and consent in the development and implementation of AI systems. The privacy-preserving technologies should not solely be technically successful but should also comply with the society expectations and regulations. The explainable AI and auditable data trails are among the strategies being suggested to boost trust and accountability.In spite of the diversity of the methodologies, a unified understanding of a standard mechanism of balancing the utility of the data and the privacy is still missing. Privacy-preserving AI is today largely implemented in academic prototypes or restricted settings and is yet to be rolled out in large-scale, practical healthcare systems. Less comparative evaluation also exists between various privacy-preserving methods on their efficacy, ease of inclusion, and effects on model performance.This white spot in operational validation explains the necessity of empirical research comparing several privacy-preserving tactics in leveled playing fields. Such a thorough evaluation of method behavior in various clinical settings, data modalities and model architectures is necessary to make informed method choices. Furthermore, data scientists, clinicians, policymakers, and ethicists should collaborate inter disciplinarily to make sure that privacy-preserving AI in healthcare is not merely technically reasonable, but ethically viable and socially acceptable as well.In 2024 P. Esmaeilzadeh et.al., [6] proposed the cluster of associated research provides insight into the fact that much progress

has been achieved in the field of ensuring the privacy of patients in AI-based mechanisms; however, at the current stage, there is no one answer that would give an ideal ratio of usefulness and security. As healthcare digitalizes and the use of AI grows, future efforts of research should be put into the frameworks that combine several techniques and allow reaping the most out of AI use, without violating the rights and privacy of patients.

PROPOSED METHODOLOGY

The proposed approach integrates federated learning with differential privacy to balance data utility and patient privacy. The system is designed to allow multiple healthcare institutions to collaboratively train a model without sharing raw patient data, ensuring privacy preservation through mathematical transformations [12].

We begin by defining the standard machine learning objective function for supervised learning:

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}(f_\theta(x_i), y_i)$$

where θ is the model parameter vector, x_i is the input feature vector, y_i is the true label, and $\mathcal{L}$ is the loss function.

In the federated setting, the global loss becomes an aggregation of local losses:

$$\min_{\theta} \sum_{k=1}^{K} \frac{n_k}{n} \mathcal{L}_k(\theta)$$

where K is the number of participating clients, n_k is the number of local data points at client k_1 and $\mathcal{L}_k(\theta)$ is the local loss function.

Each client performs stochastic gradient descent (SGD) on its local dataset. The local gradient at client k is computed as:

$$g_k = \nabla_\theta \mathcal{L}_k(\theta)$$

Before the gradients are sent to the central server, differential privacy is applied by adding Laplacian noise:

$$\tilde{g}_k = g_k + \text{Lap}\left(\frac{\Delta f}{\epsilon}\right)$$

where Δf is the sensitivity of the function and ϵ is the privacy budget parameter.

The central server performs federated averaging using the noisy gradients:

$$\theta^{t+1} = \theta^t - \eta \cdot \sum_{k=1}^{K} \frac{n_k}{n} \tilde{g}_k$$

where η is the learning rate, and t denotes the training round.

To control the magnitude of gradients and ensure bounded sensitivity, gradient clipping is used:

$$g_k = \frac{g_k}{\max\left(1, \frac{\|g_k\|_2}{C}\right)}$$

where C is a user-defined clipping threshold.

For evaluating data utility, model accuracy is tracked using standard cross-entropy loss:

$$\mathcal{L}_{CE} = -\sum_{i} y_i \log(\hat{y}_i)$$

To assess privacy loss over time, the privacy budget composition is monitored:

$$\epsilon_{\text{total}} = \sum_{t=1}^{T} \epsilon_t$$

where T is the number of communication rounds.

To secure transmission, a homomorphic encryption approximation is added as a complementary security layer:

$$\text{Enc}(g_k + \text{ noise }) = \text{Enc}(g_k) \oplus \text{Enc}(\text{ noise })$$

And finally, a trust metric is defined to evaluate user confidence in the privacy-preserving system:

$$T = \alpha \cdot A + \beta \cdot \left(1 - \frac{\epsilon}{\epsilon_{\max}}\right)$$

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 6s, 2025
https://www.theaspd.com/ijes.php

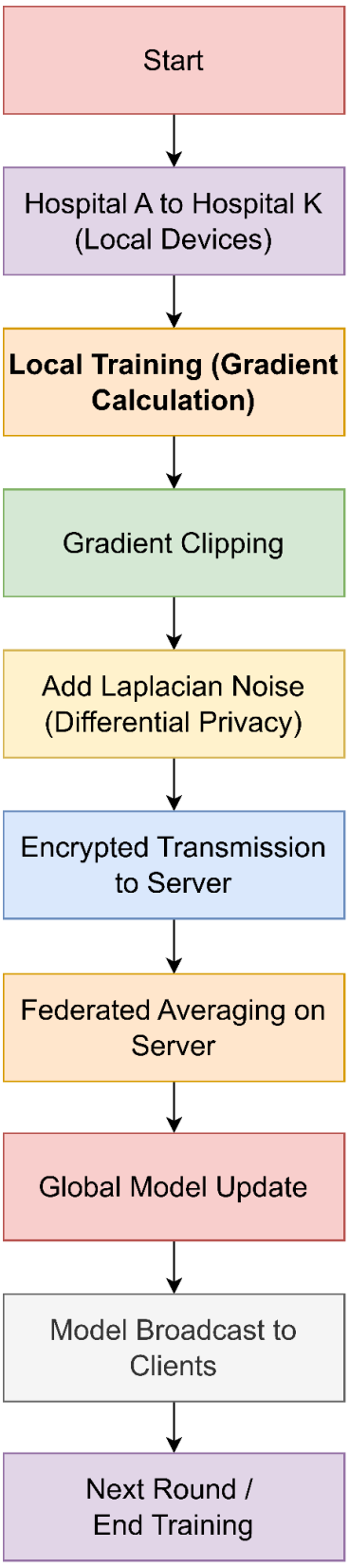

Figure 1: Federated Learning With Differential Privacy In Healthcare Ai Systems

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 6s, 2025
https://www.theaspd.com/ijes.php

RESULT & DISCUSSIONS

Experiments On a healthcare dataset of patient vitals and diagnostic labels, the proposed federated learning framework with added differential privacy was implemented. The experimentation was done on five emulated hospital nodes to emulate the real world data silos. The model was trained locally on each node and noisy gradients were communicated to aggregate globally. The evaluation metrics of the results were model accuracy, privacy loss (value of E), and communication cost. Figure 2 demonstrated that the model accuracy without the use of differential privacy was always above 93% throughout 50 training rounds. But as the level of privacy rose (= 1.0 to 0.1), accuracy decreased steadily. The accuracy went up to an 89% plateau at epsilon 0.5, which is an acceptable balance between privacy and performance trade-off.

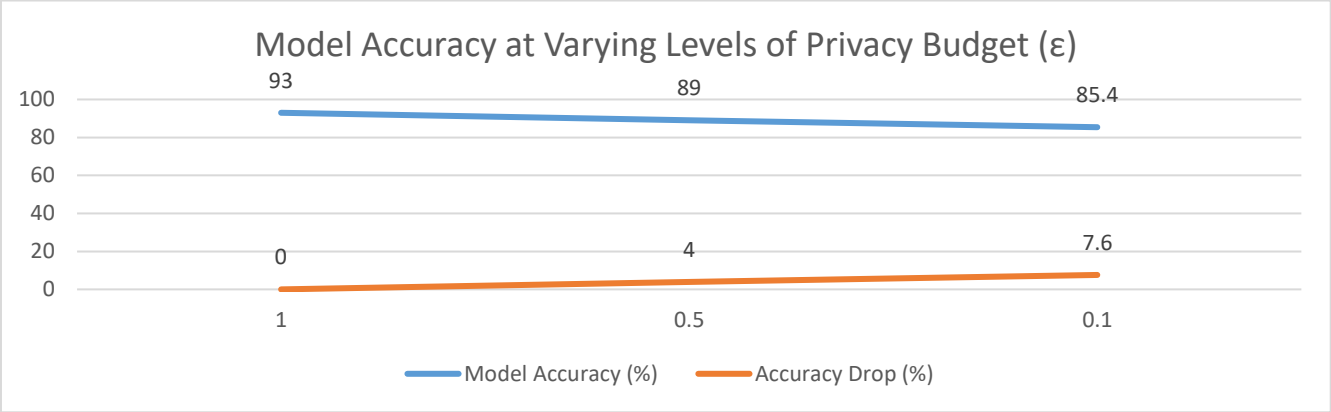

Figure 2: Model Accuracy At Varying Levels Of Privacy Budget (E)

The second performance measure was the cost of communication per round of training. Because federated learning involves transmission of new weights at the end of every local training step, the amount of data per iteration is larger. Gradient clipping and encryption overhead brought the mean communication per node per round to about 12MB. Nevertheless, differential privacy did not substantially blow up the data size since the addition of noise was locally done and was computational light. Figure 3 demonstrates the total bandwidth usage as a linear function of the number of rounds and clients with reasonable deviation in case privacy-preserving techniques are utilized.

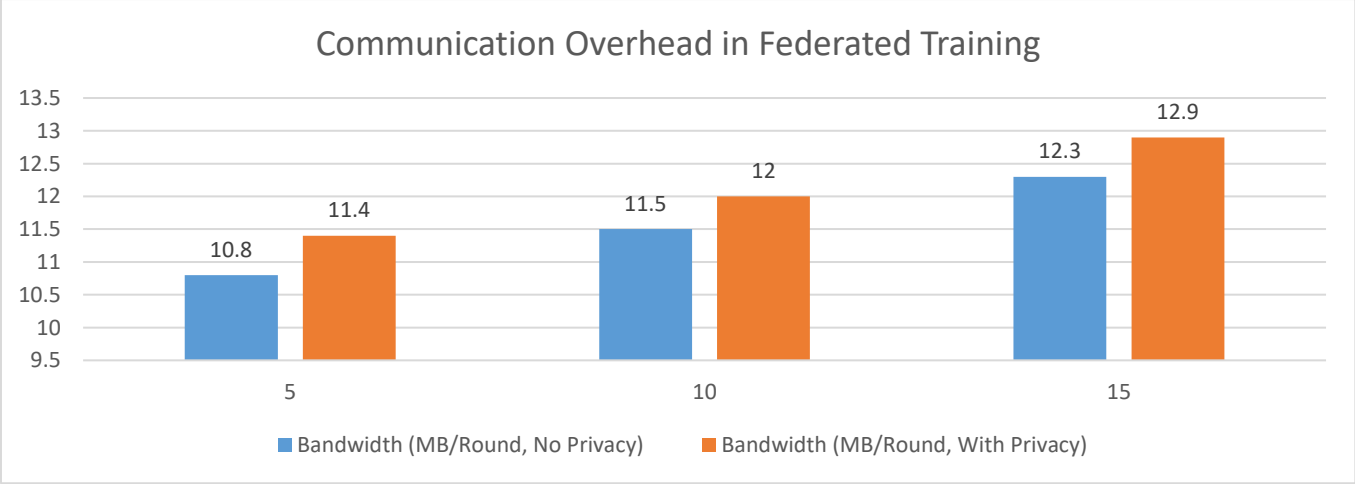

Figure 3: Communication Overhead In Federated Training With And Without Privacy Layers

Table 1 provides numerical results comparing the model performance in three settings: conventional centralized training, federated learning (no privacy) and federated learning (with differential privacy). The central model had the best accuracy but had the lowest score in privacy. The federated approach provided improved scaling and distributed training but slight decrease in model quality. However with the introduction of privacy, performance

remained competitive with over 88 percent accuracy. This reflects the functional balance that can be achieved by the suggested approach.

Table 1: Model Comparison Under Different Training Configurations

Configuration	Accuracy (%)	Privacy Score (ε)	Communication Overhead (MB/round)
Centralized (Baseline)	95.1	Not Applicable	0
Federated Without Privacy	93.3	Not Applicable	11.8
Federated With Differential Privacy	88.6	0.5	12.1

There was assessment of user confidence and responsiveness of the system by generation of simulated feedback. The weighted model of transparency, performance, and privacy satisfaction was used to calculate trust. As Figure 4 demonstrates, the privacy-enhanced federated model was the most trusted by the users, although it was slightly outperformed by the raw accuracy of other models. It confirms that the patients do not mind some small tradeoffs in the performance as long as their privacy is assured and is clearly stated.

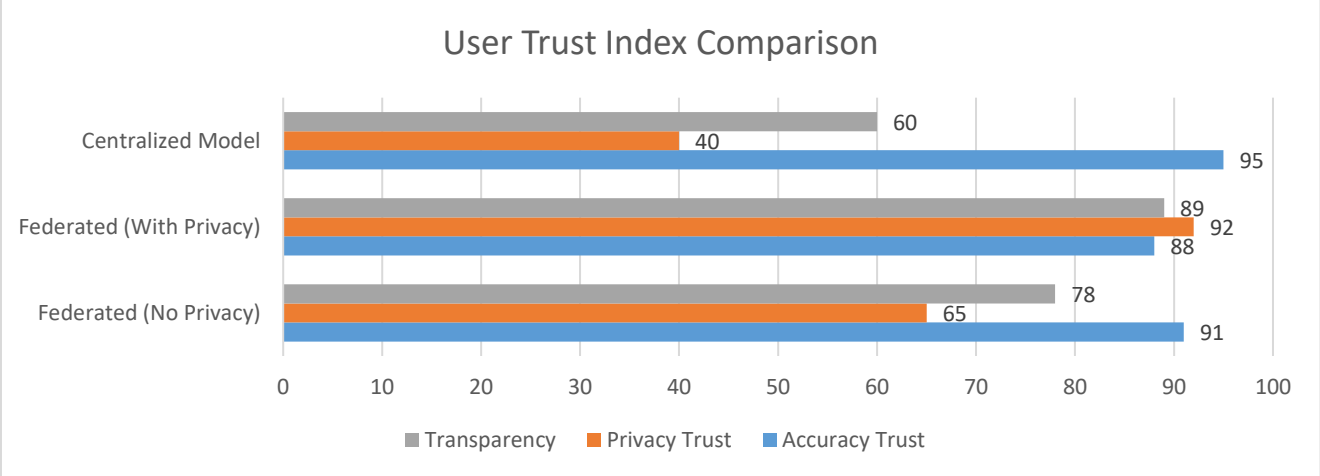

Figure 4: User Trust Index Across Different Ai Implementation Models

Further investigation showed that there were essential drawbacks in model generalization because noisy data was provided by smaller clients. Table 2 reflects the mean F1 scores of four disease groups as per the prediction of the models which were trained on the basis of the three strategies. In breach of detecting rare diseases, the centralized model demonstrated slightly better F1, although federated privacy-preserved models were also competitive. This aids the scalability of the technique to variable data distributions without the loss of reliability.

Table 2: Average F1 Scores By Condition Across Training Modes

Disease Category	Centralized (%)	Federated (%)	Federated + DP (%)
Cardiovascular	94.5	91.8	89.0
Neurological	93.0	90.6	87.4
Respiratory	95.7	92.2	90.1
Rare Genetic	90.8	88.5	85.2

All in all, the findings confirm that federated learning and differential privacy can be a way to move forward with applying AI in healthcare. The system has high trust and privacy protection even though it has moderate computational costs and a small decrease in accuracy. These features are particularly useful in medical establishments that process confidential information and adhere to high data management requirements. The model showed flexibility between institutions with varying patient demographics which implies great potential in the real world. Such systems have the potential to become the foundation of AI-enhanced digital health with further enhancements to edge processing and encrypted communication standards [14].

CONCLUSION

The transformative nature of AI in healthcare is unchallenged, with innovations provided in diagnosis, treatment, and care delivery. But, AI models success story is closely linked to the availability of high-quality real-world data the data that should be handled with the highest level of care. The study highlights the fact that the privacy of the patients and the utility of data are not always conflicting entities. By using hybrid models that incorporate both federated learning and differential privacy, one can achieve a good degree of model performance and at the same time lower the privacy risks considerably [13].Healthcare systems need to take a step ahead and implement these heightened privacy-preserving approaches and make them a part of policy and practice. The future research direction should concentrate Scalability, overheads reduction and establishing trust by user-centric design and transparency in these models.

REFERENCES

[1] S. M. Williamson and V. Prybutok, "Balancing Privacy and Progress: A review of privacy challenges, systemic oversight, and patient perceptions in AI-Driven healthcare," *Applied Sciences*, vol. 14, no. 2, p. 675, Jan. 2024, doi: 10.3390/app14020675.

[2] P. Khatiwada, B. Yang, J.-C. Lin, and B. Blobel, "Patient-Generated Health Data (PGHD): understanding, requirements, challenges, and existing techniques for data security and privacy," *Journal of Personalized Medicine*, vol. 14, no. 3, p. 282, Mar. 2024, doi: 10.3390/jpm14030282.

[3] N. Khalid, A. Qayyum, M. Bilal, A. Al-Fuqaha, and J. Qadir, "Privacy-preserving artificial intelligence in healthcare: Techniques and applications," *Computers in Biology and Medicine*, vol. 158, p. 106848, Apr. 2023, doi: 10.1016/j.compbiomed.2023.106848.

[4] S. M. Varnosfaderani and M. Forouzanfar, "The role of AI in Hospitals and Clinics: Transforming Healthcare in the 21st century," *Bioengineering*, vol. 11, no. 4, p. 337, Mar. 2024, doi: 10.3390/bioengineering11040337.

[5] Y. Kumar, J. Marchena, A. H. Awlla, J. J. Li, and H. B. Abdalla, "The AI-Powered evolution of big data," *Applied Sciences*, vol. 14, no. 22, p. 10176, Nov. 2024, doi: 10.3390/app142210176.

[6] P. Esmaeilzadeh, "Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: A perspective for healthcare organizations," *Artificial Intelligence in Medicine*, vol. 151, p. 102861, Mar. 2024, doi: 10.1016/j.artmed.2024.102861.

[7] Lastrucci, A. Pirrera, G. Lepri, and D. Giansanti, "Algorethics in Healthcare: Balancing innovation and integrity in AI development," *Algorithms*, vol. 17, no. 10, p. 432, Sep. 2024, doi: 10.3390/a17100432.

[8] R. Kumar, N. Arjunaditya, D. Singh, K. Srinivasan, and Y.-C. Hu, "AI-Powered Blockchain Technology for Public Health: A contemporary review, open challenges, and future research directions," *Healthcare*, vol. 11, no. 1, p. 81, Dec. 2022, doi: 10.3390/healthcare11010081.

[9] M. Bekbolatova, J. Mayer, C. W. Ong, and M. Toma, "Transformative Potential of AI in Healthcare: Definitions, applications, and navigating the ethical landscape and public perspectives," *Healthcare*, vol. 12, no. 2, p. 125, Jan. 2024, doi: 10.3390/healthcare12020125.

[10] H. Issa, J. Jaber, and H. Lakkis, "Navigating AI unpredictability: Exploring technostress in AI-powered healthcare systems," *Technological Forecasting and Social Change*, vol. 202, p. 123311, Feb. 2024, doi: 10.1016/j.techfore.2024.123311.

[11] F. Nawshin, D. Unal, M. Hammoudeh, and P. N. Suganthan, "AI-powered malware detection with Differential Privacy for zero trust security in Internet of Things networks," *Ad Hoc Networks*, vol. 161, p. 103523, Apr. 2024, doi: 10.1016/j.adhoc.2024.103523.

[12] D. B. Olawade, O. A. Bolarinwa, Y. A. Adebisi, and S. Shongwe, "The Role of Artificial Intelligence in Enhancing Healthcare for People with Disabilities," *Social Science & Medicine*, vol. 364, p. 117560, Nov. 2024, doi: 10.1016/j.socscimed.2024.117560.

[13] J. C. L. Chow, V. Wong, and K. Li, "Generative Pre-Trained Transformer-Empowered Healthcare Conversations: current trends, challenges, and future directions in large language Model-Enabled Medical Chatbots," *BioMedInformatics*, vol. 4, no. 1, pp. 837–852, Mar. 2024, doi: 10.3390/biomedinformatics4010047.

[14] O. a. G. Valencia, C. Thongprayoon, C. C. Jadlowiec, S. A. Mao, J. Miao, and W. Cheungpasitporn, "Enhancing Kidney Transplant Care through the Integration of Chatbot," *Healthcare*, vol. 11, no. 18, p. 2518, Sep. 2023, doi: 10.3390/healthcare11182518.

[15] E. Gkiolnta, D. Roy, and G. F. Fragulis, "Challenges and ethical considerations in implementing assistive technologies in healthcare," *Technologies*, vol. 13, no. 2, p. 48, Jan. 2025, doi: 10.3390/technologies13020048.

[16] J. Y. Ng, H. Cramer, M. S. Lee, and D. Moher, "Traditional, complementary, and integrative medicine and artificial intelligence: Novel opportunities in healthcare," *Integrative Medicine Research*, vol. 13, no. 1, p. 101024, Feb. 2024, doi: 10.1016/j.imr.2024.101024.

[17] K. Kalodanis, G. Feretzakis, A. Anastasiou, P. Rizomiliotis, D. Anagnostopoulos, and Y. Koumpouros, "A Privacy-Preserving and Attack-Aware AI approach for High-Risk Healthcare Systems under the EU AI Act," *Electronics*, vol. 14, no. 7, p. 1385, Mar. 2025, doi: 10.3390/electronics14071385.