Inclusive, Differentially Private Federated Learning for Clinical Data

Santhosh Parampottupadam[1,2][0009-0009-9401-887X], Melih Coşğun[3][0009-0008-3596-8376], Sarthak Pati[5][0000-0003-2243-8487], Maximilian Zenk[1,2][0000-0002-8933-5995], Saikat Roy[1][0000-0002-0809-6524], Dimitrios Bounias[1,2][0000-0002-3361-1698], Benjamin Hamm[1,2][0009-0003-4818-8700], Sinem Sav[3][0000-0001-9096-8768], Ralf Floca[1][0000-0003-3218-3377], and Klaus Maier-Hein[1,2,4][0000-0002-6626-2463]

[1] German Cancer Research Center (DKFZ), Heidelberg, Division of Medical Image Computing, Germany
[2] Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany
[3] Department of Computer Engineering, Bilkent University
[4] School of Medicine, Indiana University Pattern Analysis and Learning Group, Department of Radiation Oncology, Heidelberg University Hospital, 69120 Heidelberg, Germany
[5] Medical Research Group, MLCommons, San Francisco, CA, USA

Abstract. Federated Learning (FL) offers a promising approach for training clinical AI models without centralizing sensitive patient data, yet its real-world adoption is hindered by challenges in privacy, resource constraints, and compliance. Existing differential privacy (DP) approaches often apply uniform noise, which disproportionately degrades model performance even among well-compliant institutions. In this work, we propose a novel compliance-aware FL framework that enhances DP by adaptively adjusting noise based on quantifiable client compliance scores. Additionally, we introduce a compliance scoring tool based on key healthcare and security standards to promote secure, inclusive, and equitable participation across diverse clinical settings. Extensive experiments on the public datasets demonstrate that integrating under-resourced, less compliant clinics with highly regulated institutions yields accuracy improvements of up to 15% over traditional FL. This work advances FL by balancing privacy, compliance, and performance, making it a viable solution for real-world clinical workflows in global healthcare.

Keywords: Compliance-Aware Clinical Federated Learning · Privacy-Preserving FL · Adaptive Compliance· Resource-Efficient DP.

1 Introduction

Artificial Intelligence (AI) can advance healthcare through improved diagnostics and personalized treatments, but privacy concerns and regulatory constraints limit its adoption. Federated Learning (FL) [22] enables decentralized model training, preserving data privacy and security while supporting collaborative clinical AI development. Despite its potential, FL in healthcare [30] faces challenges in data security, privacy, and inclusivity. FL systems are vulnerable to reconstruction attacks, where model updates can reveal sensitive information [8,32]. Differential privacy (DP) has been integrated into FL to mitigate these risks, providing theoretical guarantees against data reconstruction and inference attacks [2,10]. However, DP introduces trade-offs by adding noise to model updates, often degrading performance [3]. Traditional DP methods apply noise uniformly across clients [21], overlooking disparities such as compliance, resources [28,23].

Healthcare FL faces significant challenges due to institutional heterogeneity, with DP imposing high computational demands that often require specialized hardware [6]. Clinical sites with lower patient loads struggle to participate due to resource constraints, compliance gaps, and coordination overhead [26,9,19]. Real-world FL studies [25,29] demonstrate feasibility but rely on trust-based federations, marginalizing smaller institutions. Balancing privacy and utility in DP requires clear trade-offs, as any DP implementation impacts model performance. A review of 612 studies found only 5.2% involved real-world clinical applications, highlighting the need for FL frameworks that ensure privacy, inclusivity, and equitable participation while addressing compliance and computational barriers [19,5,6].

This paper proposes a novel compliance-aware FL framework to enhance privacy in healthcare by dynamically integrating DP with client compliance scores. The framework introduces a customizable compliance scoring tool aligned with key healthcare standards to ensure privacy, security, and interoperability while maintaining inclusivity. It incorporates privacy concepts from various regulatory and best-practice frameworks such as patient consent management [15], anonymization practices [13,17], audit logs & network security [12], data encryption & secure infrastructure [24], ethical AI policies [1], interoperability [16], and data & model training quality. These standards collectively address privacy risks, enforce secure data handling, and promote equitable FL scalability in clinical environments.

To mitigate manipulation risks in untrusted client settings, our framework performs adaptive server-side DP, optimizing noise injection to balance privacy and utility [31]. By adapting noise levels to client compliance scores, it ensures robust performance in resource-constrained healthcare environments. The compliance scoring tool enables investigators to weigh regulatory adherence, data integrity, and security protocols, fostering tailored and trustworthy FL deployments. We evaluated our method on multiple public datasets [33] and aggregation methods [22,20,27], and quantified overall accuracy gains of 1% to 15%.

This manuscript's contributions are: *i)* a compliance-aware FL framework with adaptive DP, adjusting noise based on client compliance to enhance fairness and inclusivity, *ii)* a web-based compliance scoring tool aligned with healthcare and security standards to provide quantifiable compliance scores, and *iii)* implementation of adaptive server-side DP, enabling resource-constrained clinics to participate while balancing privacy and performance.

2 Methods

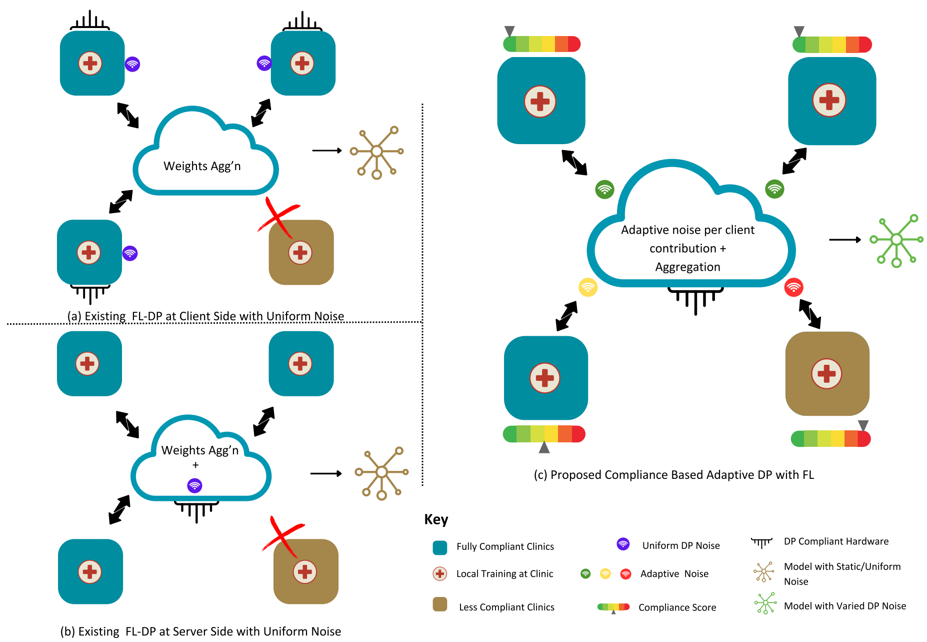

Fig. 1. (a) Existing FL with client-side DP uses uniform noise, requiring DP-compliant hardware, limiting less compliant, resource-constrained clinics. (b) Server-side DP adds uniform noise post-aggregation, reducing privacy-utility efficiency and further excluding less compliant clinics. (c) Our compliance-aware adaptive DP applies per-client noise before aggregation, enabling participation from low-resource, less compliant clinics while optimizing privacy and performance.

Compliance Scoring Mechanism. Our compliance scoring tool enables experiment organizers to assign weights to various factors (see Table 2 for an example) and configure corresponding options, offering flexible, customized evaluation for diverse clinical settings. The overall compliance score (S_c) for each client is determined by assessing all the factors and is calculated as follows:

$$S_c = \frac{\sum_{i=1}^{n} (w_i \cdot s_i)}{\sum_{i=1}^{n} w_i} \tag{1}$$

where n is the total number of compliance factors, w_i is the weight assigned to factor i, and s_i is the selected option score for factor i. For instance, the *anonymization practices* factor offers three options: *ISO/TS 25237:2017 Fully Anonymized* (Score 1.0), *Pseudonymized (Partial Anonymization)* (Score 0.7), and *No Anonymization* (Score 0.5), with the tool defaulting to a 0.5 threshold, adjustable by experiment owners, including setting it to 0 if needed.

Algorithm 1 Adaptive Noise-Based Differential Privacy in Federated Learning

1: Initialize $GLOBAL\_MODEL$
2: **for** round = 1 to FED_ROUNDS **do**
3:     **Client Training:**
4:     **for** each client i **do**
5:         $CLIENT_i \leftarrow \textsc{Copy}(GLOBAL\_MODEL)$
6:         $CLIENT_i \leftarrow \textsc{Train}(CLIENT_i, data_i, epochs = 1)$
7:     **end for**
8:     Send $\{CLIENT_i\}$ to aggregator
9:     **DP Processing:**
10:     **for** each client i **do**
11:         $DP_i \leftarrow \textsc{Copy}(CLIENT_i)$
12:         $DP_i \leftarrow \text{DP}\textsc{Train}(DP_i, agg\_data, \eta = \textsc{AdaptiveNoise}(c_i))$
13:     **end for**
14:     **Aggregation:**
15:     $GLOBAL\_MODEL \leftarrow \textsc{FedAvg}(\{DP_i\})$ ▷ Fed Median/Prox/Yogi/Adam
16:     **Broadcast** $GLOBAL\_MODEL$ to clients
17: **end for**
18: **return** $GLOBAL\_MODEL$

Noise Multiplier Calculation. To implement DP adaptively, noise levels are dynamically adjusted based on client compliance scores. The noise multiplier (N_m) is computed as: $N_m = (1.0 - S_c) + \text{Min Noise Multiplier}$, where S_c denotes the client's compliance score, and Min Noise Multiplier (set to $1e$-10 in this experiment) ensures baseline privacy. This approach ensures that clients with lower compliance scores need higher noise levels. Noise can be tuned or clipped per FL aggregation strategy, protecting data while preserving model quality and ensuring secure FL participation.

Experimental Setup. Experiments were conducted with a batch size of 32, 50 FL training rounds, a learning rate of 0.001, and images resized to 128×128. Each FL round included 3 local epochs per client, followed by 1 epoch on the aggregator dataset (at the server) using noise-injected client updates before global aggregation. This allows the model to adapt to perturbed updates, improving stability and convergence (see Algo 18). A total of 61 experiments (Table 3) were performed, including an additional data quality experiment 2. The dataset was split into 16 client subsets, with one for aggregator training with DP and another for global evaluation. Vanilla FL used the same FL rounds and learning rate but excluded DP and compliance.

Data Quality Experiment To simulate a realistic scenario and assess the "data quality" compliance factor, we degraded data for 12 clients by randomly cropping, resizing (80–100% of the original size), adding Gaussian noise ($\sigma = 0.05$), and reducing contrast to 80%. These clients received a compliance score of 0.3, while 4 trusted clients retained a score of 1.0. Compared to Experiment 4 (only 4 trusted clients), this setup showed that incorporating lower-quality data, despite its lower compliance score, can still enhance overall model performance.

Table 1. Client participation per experiment, compliant/non-compliant clients, DP settings. Non-compliant clients have compliance levels between **0.1** and **0.6**. Experiment 1 includes **12** non-compliant clients, split into two groups of **6**, each with compliance levels between **0.1** and **0.6**. Experiment 2 has **6** non-compliant clients with the same compliance range. Exp. 1-4: individual compliance-based DP. Exp. 6: DP with uniform noise post-aggregation. Baseline noise is $\mathbf{1e^{-10}}$.

Client Type	**Exp. 1**	**Exp. 2**	**Exp. 3**	**Exp. 4**	**Exp. 5**	**Exp. 6**
Compliant Clients	4	10	16	4	16-Vanilla	16
Non-Compliant Clients	12 clients	6 clients	None	None	None	None
Compliance Applied?	Yes	Yes	Yes	No	No	Yes
Minimum DP Applied?	Yes	Yes	Yes	Yes	No	Uniform DP

Implementation Details. The framework was implemented using Lightning [11], Flower [4], and ResNet-18 [14], and tested on an NVIDIA Tesla T4 GPU (16GB), demonstrating its feasibility in resource-constrained clinical settings. Compliance scores for each client were pre-assigned using a customizable web-based compliance scoring tool, simulating the role of a Principal Investigator (PI)(Table 2). This tool, grounded in established healthcare and security standards, evaluated clients on 12 compliance factors with predefined options and weights (Equation 1). These scores determined the level of noise dynamically added to client contributions2, ensuring baseline privacy with a minimum noise threshold applied across all clients. FL training began with the global model distributed to clients, who performed 3 epochs of training without DP on local datasets. The client contributions were then sent to the server, where noise proportional to compliance scores was added to each contribution. Before global aggregation, the server trained for one epoch on the noise-adjusted data using the aggregator dataset with DP [9]. The final aggregated model weights were computed using the selected FL strategy and redistributed to all clients. This iterative process was repeated for 50 FL training rounds, ensuring adaptive DP noise, robust aggregation, and inclusivity across clients with varying compliance levels. DP was integrated using Opacus [34], with minimum noise level tested ($1e-10$). Noise distribution followed the compliance score distribution, where high-compliance clients received minimal noise to preserve model performance, while low-compliance clients had higher noise applied to maintain privacy.

3 Results

Table 1 summarizes six experimental configurations on two datasets PneumoniaMNIST and BreastMNIST using various FL strategies. In these experiments, compliance-aware DP was compared against Vanilla FL across 50 experimen-

Table 2. Compliance factors and standards are customizable to fit study requirements.

Compliance Factor	Standards/Options
Data Encryption Standards	AES-256 (NIST), AES-128 (Healthcare Minimum)
Ethical AI Policies	EU AI Act, FDA Guidelines
Privacy Regulations	HIPAA, GDPR
Data Quality	DICOM Standard, Partially Validated Data
Anonymization Practices	ISO/TS 25237:2017, Pseudonymization
Interoperability Standards	HL7/FHIR Standards
Secure Network Infrastructure	NIST Cybersecurity Framework
Authentication and Authorization	MFA, RBAC
Audit Logs	SOC 2 Type II Certification
Patient Consent Management	HL7 CDA Compliant
Trusted Execution Environments	Intel SGX, AMD SEV
Local Model Training Quality	High Accuracy (>95%), Moderate Accuracy (85–95%)

tal settings (see Table 3), with different combinations of compliant and non-compliant client groups. For both datasets—PneumoniaMNIST and BreastMNIST—FedYogi achieved the highest accuracy in Experiment 1 (86.62% and 75.50%, respectively), FedAdam in Experiment 2 (85.55% and 71.49%), and FedAvg in Experiment 3 (85.64% and 73.68%). In Experiment 4 (compliant clients only), FedAvg performed best (81.28% and 65.85%). In the Vanilla FL configuration (Experiment 5), FedAdam achieved the highest accuracy for PneumoniaMNIST (86.96%), while FedYogi led for BreastMNIST (78.50%). The official AUC and ACC for PneumoniaMNIST (centralized training) are 95.6 and 86.40. For BreastMNIST, they are 89.10 and 83.30, respectively.

In addition to the experiments in Table 3, we conducted a Data Quality experiment and a realistic data quality-based compliance score experiment (see 2). The global model was evaluated on the test set using accuracy, with results across different FL strategies as follows: `dp_FedAvg` achieved 72.68%, `dp_FedYogi` 71.62%, `dp_FedAdam` 69.55%, `dp_FedMedian` 66.23%, and `dp_FedProx` 64.04%.

4 Discussion

In this manuscript, we have developed a novel compliance-aware FL framework which optimizes the privacy-utility trade-off by dynamically adjusting DP noise based on client compliance scores. We evaluated our method across multiple experiments using various aggregation strategies (FedAvg, FedProx, FedMedian, FedAdam, and FedYogi) and public datasets (PneumoniaMNIST and BreastMNIST). **Notably,** The experiment with 4 highly compliant and 12 less-compliant clients beat the 4 highly compliant-only setup, gaining 1%–15% accuracy across strategies, outperforming uniform server DP as well. This highlights that incorporating lower-compliance clients can enhance overall model performance. However, FedMedian exhibited sensitivity to compliance distribution.

Considering the experimental design (Section 2), in Experiment 1 (75% low-compliance clients), FedMedian achieved only 70.12% accuracy on PneumoniaMNIST and 50.01% on BreastMNIST (see Table 3), likely due to the median selection favoring noisy updates. In contrast, Experiment 2 (37% low-compliance clients) saw improved FedMedian accuracy (82.94% and 70.86%, respectively), nearing Vanilla FL performance. This suggests that FedMedian's effectiveness depends on compliance distribution, making it less reliable in settings with a high proportion of low-compliance clients.

Performance gains mainly benefit the principal investigator, while high compliance institutions access diverse, real-world data, improving model generalizability. FL ethically integrates data from less-compliant or resource-constrained clinics, preserving privacy with minimal DP protection for all, regardless of compliance. In rare disease studies, this collaboration is critical. For instance, a glioblastoma study [25] across 71 sites (n=6,314) saw a 33% improvement in delineating surgically targetable tumors and a 23% gain for complete tumor extent, demonstrating how high-compliance institutions benefit from the inclusion of less regulated clinics (Asia, South America, Australia) by accessing rare and geographically diverse data that would otherwise be unavailable.

We have presented a compliance-aware DP framework in FL which promotes inclusivity and reducing resource constraints without specialized hardware. While DP offers theoretical privacy guarantees [9,26], it remains the most practical alternative to trusted execution environments (hardware-dependent) and homomorphic encryption (computationally intensive). Our method minimizes computational burdens on resource-limited clinics, enabling broader participation without enforcing DP-compliant hardware [9,6]. The compliance scoring tool allows experiment administrators to customize compliance factors, aligning with global healthcare standards [18,7] to foster secure, equitable FL participation. Unlike traditional server-side DP (See Exp.6 3), which applies uniform noise across all clients, our adaptive DP mechanism adjusts noise based on compliance scores, ensuring a balanced trade-off between privacy and utility. This effectively simulates client-side DP at the server level, allowing resource-constrained clinics to contribute without requiring DP-compliant infrastructure.

5 Limitations and Future Works

While our compliance-aware FL framework advances privacy, inclusivity, and performance, some limitations remain. One is the initial trust assumption, where first-round client updates lack DP, posing a minor risk if the server is curious. Later updates mitigate this with DP, but adding minimal noise in the first round or using secure multi-party computation (SMPC) could enhance security. Additionally, the framework assumes accurate and honest compliance scores, which may not always hold. Future work could explore dynamic validation to ensure real-time compliance verification.

This work brings "privacy" closer to clinical practice by validating the framework in controlled settings with defined resource constraints and compliance parameters. Expanding its evaluation to real-world clinical environments with diverse datasets and infrastructures will provide deeper insights into its scalability and robustness. Our approach separates privacy from hardware limits, enabling resource-constrained clinics to join a more inclusive FL ecosystem. Future work could refine adaptive aggregation by compliance, balance efficiency and privacy, boost global clinical FL use, and prevent inference attacks from untrusted clients.

Table 3. Results for all combinations of Compliant Clients, Strategies, and Minimum DP Noise. Batch size is fixed at 32, and FL rounds are set to 50. Irrespective of compliance, a baseline noise of $1e-10$ is added to each model. Results for vanilla FL (no compliance, no DP) are included as a separate block. Detailed Experiment configurations are provided in Table 1.

Experiment	**Strategy**	**PneumoniaMNIST**				**BreastMNIST**			
		Acc.	**Prec.**	**Rec.**	**F1**	**Acc.**	**Prec.**	**Rec.**	**F1**
1	FedAvg	82.43	89.39	82.43	84.30	66.98	84.40	66.98	69.69
	FedMedian	70.12	81.38	70.12	71.16	50.01	36.53	50.01	42.22
	FedYogi	**86.62**	91.68	86.62	88.26	**75.50**	81.51	75.70	77.64
	FedProx	84.01	89.93	84.01	85.76	71.61	81.60	71.11	74.34
	FedAdam	84.01	89.93	84.01	85.76	64.16	60.26	64.86	66.11
2	FedAvg	85.29	90.57	85.29	86.95	70.73	78.51	70.74	73.01
	FedMedian	82.94	89.39	82.94	84.75	70.86	82.65	70.86	73.76
	FedYogi	83.84	90.32	83.84	85.68	62.21	81.46	62.24	63.58
	FedProx	84.78	90.54	84.78	86.52	64.47	76.62	64.47	66.37
	FedAdam	**85.55**	91.15	85.55	87.28	**71.49**	77.93	71.49	73.56
3	FedAvg	**85.64**	90.97	85.64	87.31	**73.68**	68.65	73.68	67.29
	FedMedian	83.67	90.73	83.67	85.61	73.24	83.98	73.29	76.24
	FedYogi	84.27	90.52	84.27	86.09	66.22	86.73	66.21	68.87
	FedProx	85.04	91.13	85.04	86.85	71.36	75.20	71.40	72.79
	FedAdam	82.99	89.91	82.99	84.87	62.97	79.36	62.97	64.58
Impact of Experiment 1 with only Compliant Clients:									
4	FedAvg	**81.28**	89.10	81.28	83.21	**65.85**	71.79	65.85	67.43
	FedMedian	79.44	87.96	79.44	81.35	62.84	73.62	62.74	64.33
	FedYogi	81.06	89.00	81.06	83.00	60.90	73.30	60.80	61.87
	FedProx	78.80	87.66	78.80	80.70	63.03	68.46	63.03	64.27
	FedAdam	79.65	88.06	79.65	81.56	54.76	57.50	54.96	51.55
Vanilla FL (No compliance Score and No DP noise):									
5	FedAvg	85.42	89.80	85.42	86.88	76.37	84.29	76.37	79.03
	FedMedian	85.34	89.96	85.34	86.85	75.81	79.79	75.81	77.38
	FedYogi	84.61	90.93	84.61	86.45	**78.50**	79.91	78.53	79.15
	FedProx	86.88	91.18	81.28	88.35	73.43	78.27	73.45	75.19
	FedAdam	**86.96**	90.10	87.00	88.12	75.18	77.89	83.65	75.18
DP with uniform noise post-weight aggregation:									
6	FedAvg	75.89	87.66	75.89	77.74	68.04	79.30	68.04	70.51
	FedMedian	76.45	88.24	76.45	78.36	68.55	68.98	73.55	74.07
	FedYogi	77.16	88.13	77.50	78.50	**72.10**	76.11	75.89	76.80
	FedProx	**79.53**	89.18	79.60	81.56	63.72	70.80	63.72	65.51
	FedAdam	79.12	89.10	78.30	89.12	63.45	79.90	73.01	75.30

References

1. Act, E.A.I.: EU Artificial Intelligence Act | Up-to-date developments and analyses of the EU AI Act — artificialintelligenceact.eu. https://artificialintelligenceact.eu/, [Accessed 11-01-2025]
2. Adnan, M., Kalra, S., Cresswell, J.C., Taylor, G.W., Tizhoosh, H.R.: Federated learning and differential privacy for medical image analysis. Scientific reports **12**(1), 1953 (2022)
3. Bagdasaryan, E., Shmatikov, V.: Differential privacy has disparate impact on model accuracy (2019), https://arxiv.org/abs/1905.12101
4. Beutel, D.J., Topal, T., Mathur, A., Qiu, X., Fernandez-Marques, J., Gao, Y., Sani, L., Kwing, H.L., Parcollet, T., Gusmão, P.P.d., Lane, N.D.: Flower: A friendly federated learning research framework. arXiv preprint arXiv:2007.14390 (2020)
5. Calvino, G., Peconi, C., Strafella, C., Trastulli, G., Megalizzi, D., Andreucci, S., Cascella, R., Caltagirone, C., Zampatti, S., Giardina, E.: Federated learning: Breaking down barriers in global genomic research. Genes **15**(12), 1650 (2024)
6. Cummings, R., Desfontaines, D., Evans, D., Geambasu, R., Huang, Y., Jagielski, M., Kairouz, P., Kamath, G., Oh, S., Ohrimenko, O., Papernot, N., Rogers, R., Shen, M., Song, S., Su, W., Terzis, A., Thakurta, A., Vassilvitskii, S., Wang, Y.X., Xiong, L., Yekhanin, S., Yu, D., Zhang, H., Zhang, W.: Advancing differential privacy: Where we are now and future directions for real-world deployment (2024), https://arxiv.org/abs/2304.06929
7. Dankar, F.K., El Emam, K.: Practicing differential privacy in health care: A review. Trans. Data Privacy **6**(1), 35–67 (Apr 2013)
8. Dimitrov, D.I., Balunović, M., Konstantinov, N., Vechev, M.: Data leakage in federated averaging (2022)
9. Dwork, C., Roth, A., et al.: The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science **9**(3–4), 211–407 (2014)
10. El Ouadrhiri, A., Abdelhadi, A.: Differential privacy for deep and federated learning: A survey. IEEE access **10**, 22359–22380 (2022)
11. Falcon, W., The PyTorch Lightning team: PyTorch Lightning (Mar 2019). https://doi.org/10.5281/zenodo.3828935, https://github.com/Lightning-AI/lightning
12. Force, J.T.: NIST Special Publication (SP) 800-53 Rev. 5, Security and Privacy Controls for Information Systems and Organizations — csrc.nist.gov. https://csrc.nist.gov/pubs/sp/800/53/r5/upd1/final, [Accessed 11-01-2025]
13. gdpr.eu: General Data Protection Regulation (GDPR). https://gdpr.eu/tag/gdpr/, [Accessed 11-01-2025]
14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
15. hhs.gov: The hipaa privacy rule. https://www.hhs.gov/hipaa/for-professionals/privacy/index.html, [Accessed 11-01-2025]
16. http://hl7.org/fhir: Overview - FHIR v5.0.0 — hl7.org. https://hl7.org/fhir/overview.html, [Accessed 11-01-2025]
17. iso.org: ISO 25237:2017 — iso.org. https://www.iso.org/standard/63553.html, [Accessed 11-01-2025]
18. Kaiser, J., Mueller, T., Kaissis, G.: Differential privacy in medical imaging applications. In: Trustworthy AI in Medical Imaging, pp. 411–424. Elsevier (2025)

19. Li, M., Xu, P., Hu, J., Tang, Z., Yang, G.: From challenges and pitfalls to recommendations and opportunities: Implementing federated learning in healthcare. arXiv preprint arXiv:2409.09727 (2024)
20. Li, T., Sahu, A.K., Zaheer, M., Sanjabi, M., Talwalkar, A., Smith, V.: Federated optimization in heterogeneous networks. Proceedings of Machine learning and systems **2**, 429–450 (2020)
21. Li, X., Zmigrod, R., Ma, Z., Liu, X., Zhu, X.: Fine-tuning language models with differential privacy through adaptive noise allocation (2024), `https://arxiv.org/abs/2410.02912`
22. McMahan, B., Moore, E., Ramage, D., Hampson, S., y Arcas, B.A.: Communication-efficient learning of deep networks from decentralized data. In: Artificial intelligence and statistics. pp. 1273–1282. PMLR (2017)
23. Nguyen, P., Silence, A., Darais, D., Near, J.P.: Duetsgx: Differential privacy with secure hardware. arXiv preprint arXiv:2010.10664 (2020)
24. NIST: Cybersecurity Framework — nist.gov. `https://www.nist.gov/cyberframework`, [Accessed 11-01-2025]
25. Pati, S., Baid, U., Edwards, B., Sheller, M., Wang, S.H., Reina, G.A., Foley, P., Gruzdev, A., Karkada, D., Davatzikos, C., et al.: Federated learning enables big data for rare cancer boundary detection. Nature communications **13**(1), 7346 (2022)
26. Pati, S., Kumar, S., Varma, A., Edwards, B., Lu, C., Qu, L., Wang, J.J., Lakshminarayanan, A., Wang, S.h., Sheller, M.J., et al.: Privacy preservation for federated learning in health care. Patterns **5**(7) (2024)
27. Reddi, S.J., Charles, Z., Zaheer, M., Garrett, Z., Rush, K., Konečný, J., Kumar, S., McMahan, H.B.: Adaptive federated optimization. In: International Conference on Learning Representations (2021), `https://openreview.net/forum?id=LkFG3lB13U5`
28. Ren, X., Yang, S., Zhao, C., McCann, J., Xu, Z.: Belt and braces: When federated learning meets differential privacy (2024), `https://arxiv.org/abs/2404.18814`
29. Schmidt, K., Bearce, B., Chang, K., Coombs, L., Farahani, K., Elbatel, M., Mouheb, K., Marti, R., Zhang, R., Zhang, Y., et al.: Fair evaluation of federated learning algorithms for automated breast density classification: The results of the 2022 acr-nci-nvidia federated learning challenge. Medical Image Analysis **95**, 103206 (2024)
30. Sheller, M.J., Edwards, B., Reina, G.A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D., Colen, R.R., et al.: Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. Scientific reports **10**(1), 12598 (2020)
31. Wang, H., Zhao, Q., Wu, Q., Chopra, S., Khaitan, A., Wang, H.: Global and local differential privacy for collaborative bandits. In: Proceedings of the 14th ACM Conference on Recommender Systems. pp. 150–159 (2020)
32. Wen, Y., Geiping, J., Fowl, L., Goldblum, M., Goldstein, T.: Fishing for user data in large-batch federated learning via gradient magnification (2022)
33. Yang, J., Shi, R., Wei, D., Liu, Z., Zhao, L., Ke, B., Pfister, H., Ni, B.: Medmnist v2-a large-scale lightweight benchmark for 2d and 3d biomedical image classification. Scientific Data **10**(1), 41 (2023)
34. Yousefpour, A., Shilov, I., Sablayrolles, A., Testuggine, D., Prasad, K., Malek, M., Nguyen, J., Ghosh, S., Bharadwaj, A., Zhao, J., Cormode, G., Mironov, I.: Opacus: User-friendly differential privacy library in PyTorch. arXiv preprint arXiv:2109.12298 (2021)