MANTECH Publications

Journal of Converging Technologies in Intelligent
Volume 1, Issue 2, May-August 2025

Federated Learning and Privacy-Preserving AI: Emerging Approaches for Distributed Intelligent Computing — A Systematic Review of Algorithms, Privacy Mechanisms, and Deployment Trade-offs

Sumit K. Sengupta[1], Moumita D. Pal[2], Subhajit R. Dutta[3]

ABSTRACT

The growth of sensitive data held by hospitals, banks, mobile devices, and industrial edge nodes has created a basic conflict between the demand for data-driven intelligence and the legal and ethical duty to protect individual privacy. Federated learning (FL) addresses this conflict by moving model training to the data instead of moving data to a central server, yet FL alone does not guarantee privacy, because shared gradients and model updates can leak training information. This review-based study synthesises 35 peer-reviewed papers, preprints, books, and regulatory documents published between 2009 and 2023, following a protocol adapted from systematic review guidelines for software engineering. Each source was analysed along five dimensions: statistical heterogeneity, communication cost, privacy threat model, privacy mechanism, and deployment domain. The reported evidence shows that federated averaging reduces communication rounds by one to two orders of magnitude; that label-skewed non-IID data can reduce accuracy by up to 51% on CIFAR-10; that differentially private training retains roughly 97% accuracy on MNIST at a privacy budget of $\varepsilon = 8$ but falls to about 90% at $\varepsilon = 0.5$; and that secure aggregation protects individual updates but not the aggregate model. A qualitative comparison of five privacy mechanisms is presented, and a layered defence-in-depth framework is proposed for selecting mechanisms in cross-device and cross-silo deployments.

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

KEYWORDS: *Federated learning, Privacy-preserving AI, Differential privacy, Secure aggregation, Homomorphic encryption, Non-IID data, Edge intelligence, Distributed computing, Gradient leakage, Data protection*

INTRODUCTION

Modern artificial intelligence depends on large and diverse datasets, but the most valuable data are often the most sensitive. Patient records sit inside hospital systems, transaction histories sit inside banks, and typing behaviour, location traces, and health signals sit on personal devices. Data-protection regimes such as the European General Data Protection Regulation [1] and India's Digital Personal Data Protection Act, 2023 [2] restrict how such data may be collected, transferred, and processed, and they place particular emphasis on purpose limitation and data minimisation. Conventional machine learning, which pools raw data in a central repository, is therefore increasingly difficult to justify legally, and it also creates a single point of failure for breaches.

Federated learning was proposed as an architectural answer to this difficulty. In the federated averaging (FedAvg) algorithm, a server sends the current global model to a set of clients, each client trains it on local data for several epochs, and the server forms the next global model as a data-size-weighted average of the returned models, $w(t+1) = \Sigma k\ (nk/n)\ wk(t+1)$ [3]. Raw data never leave the client. The approach has since grown into a broad research field with its own taxonomy, algorithms, and open problems [4, 5, 6]. However, keeping data local is not the same as keeping data private. Gradients can be inverted to reconstruct training images almost exactly [7, 8], trained models can reveal whether a particular record was used in training [9], and a single malicious participant can implant a backdoor in the global model [10]. Formal privacy tools, namely differential privacy [11, 12], secure aggregation [13], and homomorphic encryption [14], are consequently combined with FL to form what is now called privacy-preserving AI.

The central problem addressed in this review is the absence of a consolidated, evidence-based view of how FL algorithms and privacy mechanisms interact. Choosing a stronger privacy mechanism almost always costs accuracy, computation, or bandwidth, and these costs are amplified when client data are statistically heterogeneous and devices are resource-constrained. The practical significance is considerable: hospital consortia [15], mobile keyboards [16], and

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

edge-computing networks [17] are already deploying FL, and practitioners need guidance on which protections are necessary, which are sufficient, and which are affordable in a given setting.

LITERATURE REVIEW

Foundations and Taxonomy of Federated Learning

Yang et al. classified FL into horizontal, vertical, and federated transfer learning according to how the feature and sample spaces of participants overlap [5]. Horizontal FL, in which clients share a feature space but hold different users, is the setting of FedAvg and of most mobile applications. Vertical FL applies when organisations hold different attributes about the same individuals, for example a bank and a retailer. Kairouz et al. added the orthogonal distinction between cross-device FL, with millions of unreliable and intermittently available phones, and cross-silo FL, with a few dozen reliable institutions [4]. Table 1 summarises these settings and their consequences for privacy design, and Figure 1 shows the generic privacy-enhanced FL workflow assumed in this review.

Table 1. Taxonomy of Federated Learning Settings and Privacy Implications

Setting	Participants and data	Typical scale	Dominant privacy concern
Horizontal FL	Same features, different users (e.g., phones, branches)	10^3–10^7 clients	Gradient leakage, membership inference
Vertical FL	Different features, shared users (e.g., bank and retailer)	2–10 parties	Entity alignment and intermediate-value leakage
Federated transfer learning	Little overlap in users and features	Few parties	Leakage through shared representations
Cross-device	Unreliable, intermittent consumer devices	Very large	Dropouts, poisoning, per-user privacy
Cross-silo	Reliable institutions with large datasets	Tens of parties	Institution-level confidentiality, regulation

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

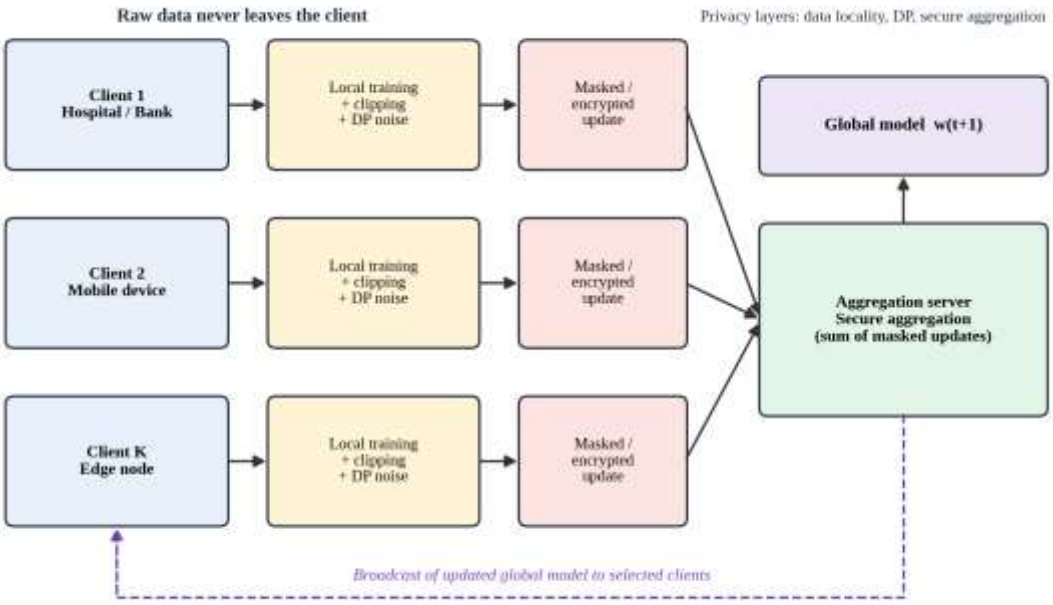


Figure 1. Generic privacy-enhanced federated learning workflow combining local clipping and differential-privacy noise, secure aggregation, and global model broadcast

Statistical and Systems Heterogeneity

Real federated data are non-IID: clients differ in label distribution, feature distribution, and data volume. Zhao et al. showed that when each client holds only one or two classes, FedAvg accuracy can fall by up to 11% on MNIST, 51% on CIFAR-10, and 55% on a keyword-spotting task, because local updates diverge from the global optimum [18]. Hsu et al. confirmed the sensitivity of visual classification to the degree of label skew [19]. Remedies include a proximal term that restricts local drift (FedProx) [20], control variates that correct client drift (SCAFFOLD) [21], and meta-learning-based personalisation that gives each client an adapted model [22]. Communication is the second systems bottleneck: structured and sketched updates can reduce uplink traffic by up to two orders of magnitude [23], and sparse ternary compression is robust to non-IID data [24]. Standard benchmarks such as LEAF now allow such methods to be compared reproducibly [25].

Privacy Threats in Federated Systems

Surveys catalogue threats across the training and inference pipeline [26, 27]. Deep leakage from gradients recovers pixel-level training images and text from shared gradients [7], and later work showed that this is feasible even for trained, deep networks with large batches [8]. Membership inference determines whether a given record belonged to the training set [9]. On

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

the integrity side, a single participant can scale its update so that it replaces the global model and embeds a backdoor [10], while Byzantine-tolerant aggregation rules such as Krum limit the influence of arbitrary faulty updates [28]. These threats show that confidentiality, output privacy, and integrity must be treated as separate requirements.

Privacy-Preserving Mechanisms

Differential privacy (DP) bounds the influence of any single record on the output through the parameters (ε, δ) [11]. DP-SGD clips per-example gradients and adds calibrated Gaussian noise, with a moments accountant that tracks cumulative privacy loss [12]; Rényi DP provides tighter composition [29]. McMahan et al. extended this to user-level DP for recurrent language models [30], and Wei et al. analysed the convergence of FL under DP noise [31]. Secure aggregation lets the server learn only the sum of client updates by using pairwise random masks that cancel on summation, and it tolerates client dropout [13]. Homomorphic encryption allows computation on ciphertexts [14], and secure multi-party computation underpins systems such as SecureML [32]. A comprehensive system-level view of these options is given by Li et al. [33].

RESEARCH GAP

Existing surveys cover FL optimisation [4], threats [26, 27], or system architecture [33], but they usually treat statistical heterogeneity, communication efficiency, and privacy protection as separate topics. Three gaps remain. First, there is no compact cross-comparison that places privacy guarantees, computational and communication overhead, and accuracy impact on a common footing. Second, the interaction between non-IID data and DP noise is rarely discussed in a way that supports design decisions. Third, little guidance exists for regulated, resource-constrained environments, such as small hospitals and cooperative banks operating under India's data-protection law [2], where cross-silo cryptographic protection may be feasible while device-scale compute is limited.

OBJECTIVES

- To classify federated learning settings and identify the privacy requirements associated with each.
- To synthesise reported evidence on communication efficiency and the effect of non-IID data on model accuracy.

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

- To compare differential privacy, secure aggregation, homomorphic encryption, secure multi-party computation, and robust aggregation in terms of protection, cost, and utility.
- To map documented attacks to the defences that address them and to expose residual risks.
- To propose a layered selection framework for privacy-preserving FL in cross-device and cross-silo deployments.

METHODOLOGY

Review Design and Source Selection

This is a review-based study. The protocol was adapted from the guidelines for systematic literature reviews in software engineering [34], with the purpose of producing a structured narrative synthesis rather than a statistical meta-analysis, because the primary studies use different datasets, models, and privacy budgets. Candidate sources were identified in IEEE Xplore, the ACM Digital Library, ScienceDirect, SpringerLink, arXiv, and the proceedings of NeurIPS, ICML, ICLR, and AISTATS, using combinations of the terms federated learning, differential privacy, secure aggregation, homomorphic encryption, gradient leakage, and non-IID. Sources were included if they introduced a foundational algorithm, reported a quantitative result on a public benchmark or real deployment, analysed a documented attack, or defined a regulatory requirement. Opinion pieces without technical content were excluded. The final corpus comprised 35 sources.

Analytical Framework

Each source was coded along five dimensions: (i) data setting (horizontal, vertical, cross-device, cross-silo), (ii) heterogeneity handling, (iii) threat model (honest-but-curious server, malicious client, external adversary), (iv) privacy mechanism and its formal guarantee, and (v) cost, expressed as accuracy change, computation, and communication. Quantitative values are quoted only as reported by the original authors and are compared only within the same benchmark, in order to avoid invalid cross-study arithmetic. For mechanisms, a qualitative five-point rating was assigned on six criteria using the reported results and the known asymptotic cost of each protocol; these ratings are the authors' judgement and are shown in Figure 3 for relative comparison only.

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

RESULTS AND FINDINGS

Communication Efficiency and Heterogeneity

The first finding concerns communication. Allowing multiple local epochs before averaging reduced the number of communication rounds needed to reach target accuracy by 10 to 100 times relative to synchronised distributed SGD on the benchmarks reported for FedAvg [3], and update compression added up to two orders of magnitude of savings in upload size [23]. The second finding concerns heterogeneity. The accuracy loss under extreme label skew is large, but not irreversible: sharing a small globally balanced dataset equal to 5% of the data improved CIFAR-10 accuracy by about 30% in the original study [18], and drift-correcting algorithms address the same problem without sharing data [20, 21]. Table 2 collects the representative quantitative outcomes reviewed.

Table 2. Representative Quantitative Outcomes Reported in the Reviewed Literature

Source	Setting / task	Technique	Reported outcome
McMahan et al. [3]	Image and language benchmarks	FedAvg	10–100× fewer communication rounds
Konečný et al. [23]	Communication-limited uplink	Structured and sketched updates	Up to two orders of magnitude less upload
Zhao et al. [18]	Label-skewed non-IID data	FedAvg baseline	Accuracy loss up to 11% (MNIST), 51% (CIFAR-10), 55% (speech)
Zhao et al. [18]	CIFAR-10 non-IID	5% globally shared data	About 30% accuracy gain
Abadi et al. [12]	MNIST	DP-SGD	About 97% accuracy at $\varepsilon = 8$
Sheller et al. [35]	Brain-tumour segmentation	Multi-institution FL	About 99% of centralised quality
Bagdasaryan et al. [10]	Backdoor attack	Model replacement	Backdoor implanted by one participant in one round

Privacy–Utility Trade-off under Differential Privacy

DP-SGD on MNIST reached approximately 97% test accuracy at $\varepsilon = 8$, about 95% at $\varepsilon = 2$, and about 90% at $\varepsilon = 0.5$, against roughly 99% without noise [12]. Figure 2 plots these values and shows the characteristic diminishing return: reducing ε from 8 to 2 costs about two

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

percentage points, whereas reducing it from 2 to 0.5 costs a further five. The cost of user-level DP shrinks as the client population grows; for next-word prediction trained across several hundred thousand users, noise had a negligible effect on accuracy [30]. This supports the intuition that DP is naturally suited to cross-device FL, where population size dilutes the noise, and is more difficult to apply in cross-silo settings with few participants [31].

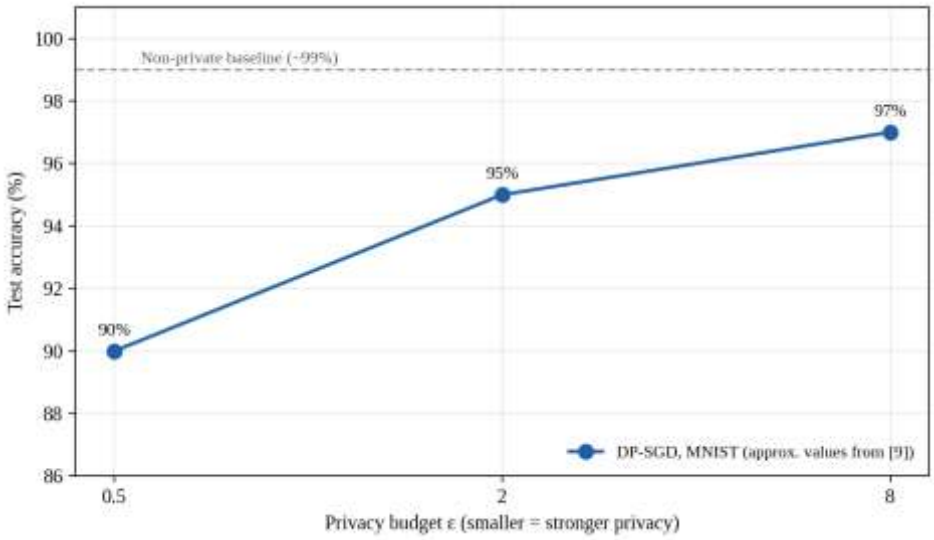


Figure 2. Privacy–utility trade-off of differentially private training on MNIST (approximate values reported in the original DP-SGD study)

Cryptographic Protection and Mechanism Comparison

Secure aggregation prevents the server from reading any individual update, but the aggregate itself is revealed, so it offers no protection against inference from the released model; combining it with DP closes that gap [13]. Its server-side computation grows quadratically with the number of participants in the original protocol, which limits scale unless clients are grouped [13]. Homomorphic encryption and secure multi-party computation offer the strongest confidentiality but impose the largest computational and communication burden [14, 32], so their natural domain is cross-silo FL with few, well-provisioned parties. Robust aggregation defends against poisoning but gives no confidentiality [10, 28]. Table 3 and Figure 3 compare the five mechanisms.

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

Table 3. Comparison of Privacy-Preserving Mechanisms for Federated Learning

Mechanism	Main threat addressed	Formal guarantee	Principal cost
Differential privacy [11, 12]	Output leakage, membership inference	(ε, δ)-DP	Accuracy loss at small ε
Secure aggregation [13]	Honest-but-curious server	Server learns only the sum	Quadratic server cost; dropout handling
Homomorphic encryption [14]	Untrusted aggregator	Computational security	High computation and ciphertext size
Secure multi-party computation [32]	Collusion among parties	Cryptographic security	High communication rounds
Robust aggregation [28]	Poisoning and Byzantine clients	Bounded influence of faulty updates	No confidentiality; accuracy loss on non-IID data

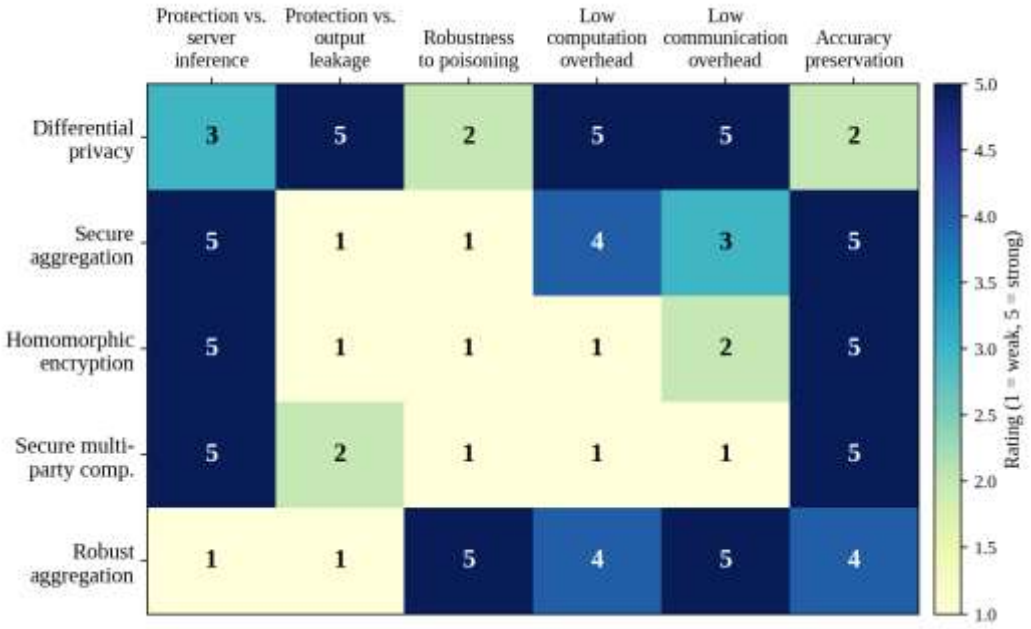


Figure 3. Qualitative comparison of privacy-preserving mechanisms across protection and cost criteria (authors' synthesis; 1 = weak, 5 = strong)

Evidence from Application Domains

In medical imaging, federated training of brain-tumour segmentation models across institutions reached about 99% of the quality of a centrally trained model while patient data stayed on site [35], and the broader digital-health literature identifies FL as an enabler of multi-institutional studies [15]. In consumer applications, FL has been used to train next-word prediction for

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

mobile keyboards [16]. In wireless and edge networks, FL reduces backhaul traffic and latency while exposing new device-level constraints [17]. Across these domains, the common observation is that the learning task itself is rarely the obstacle; the binding constraints are heterogeneity, bandwidth, and the privacy budget.

DISCUSSION

The reviewed evidence supports a defence-in-depth interpretation. No single mechanism satisfies confidentiality, output privacy, and integrity simultaneously: secure aggregation protects updates in transit, DP protects the released model, and robust aggregation protects against corrupted participants. The Figure 3 comparison shows that the protections are complementary rather than substitutable, which explains why recent system designs layer them [33]. A practical corollary is that the order of selection matters. Statistical heterogeneity should be addressed first, because noise injected into an already unstable training process compounds the accuracy loss seen in Figure 2.

For cross-device deployments, the evidence favours client-level clipping, DP noise, and secure aggregation with dropout tolerance, because large populations absorb the noise and the protocol avoids exposing single updates [13, 30]. For cross-silo deployments with a handful of institutions, such as hospital consortia, the statistical cost of DP is high, whereas the cost of homomorphic encryption or multi-party computation is tolerable given server-class hardware, so cryptographic protection supported by moderate DP is more appropriate [15, 35]. These recommendations are decision rules derived from the reviewed literature, not experimentally validated prescriptions.

The engineering significance is twofold. First, FL changes the compliance posture of an organisation: because raw records are not transferred, obligations of data minimisation and purpose limitation under the GDPR and the Indian data-protection law become easier to satisfy, although formal guarantees are still required to claim anonymisation [1, 2]. Second, practical applications extend beyond healthcare and keyboards to credit-risk modelling among cooperative banks, predictive maintenance across industrial plants, traffic and energy forecasting in smart cities, and personalised assistants on low-cost smartphones, all of which share the property that data owners will cooperate only if privacy is demonstrable.

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

LIMITATIONS

- The study is a narrative synthesis of 35 sources and does not include new experiments; quantitative results are taken from original papers whose datasets, architectures, and privacy accounting differ.
- The qualitative ratings in Figure 3 are the authors' judgement and are sensitive to assumptions about hardware, model size, and the number of participants.
- Most benchmarks reviewed are image or text tasks; evidence for tabular, time-series, and multimodal data in regulated sectors is thinner.
- Fast-moving areas such as federated learning with large foundation models are under-represented because the evidence base was limited when the corpus was assembled.

FUTURE SCOPE

- Controlled experiments that jointly vary non-IID severity, privacy budget, and client population size to quantify the interaction between heterogeneity and DP noise on the same testbed.
- Communication- and compute-efficient cryptographic protocols suitable for resource-constrained edge devices, including hardware-assisted and lattice-based approaches.
- Tighter privacy auditing methods that empirically test deployed FL systems against gradient inversion and membership inference, rather than relying only on theoretical budgets.
- Federated fine-tuning of large pre-trained models with parameter-efficient adapters, with accompanying privacy analysis.
- Domain-specific pilots in Indian healthcare and cooperative finance to evaluate regulatory alignment, governance, and incentive design for data owners.

CONCLUSION

Federated learning moves computation to the data, but privacy must still be engineered into it. This review of 35 sources shows that FedAvg-style training provides large communication savings, that non-IID data is the main threat to accuracy, and that differential privacy, secure aggregation, homomorphic encryption, multi-party computation, and robust aggregation each address a different part of the threat landscape at a different cost. Differential privacy fits cross-device settings with large populations, cryptographic protection fits cross-silo consortia, and robust aggregation is required wherever participants are not fully trusted. The proposed layered framework gives practitioners a structured starting point, and the future directions identified

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

here, particularly joint heterogeneity–privacy experiments and lightweight cryptography, define a clear agenda for research on trustworthy distributed intelligent computing.

REFERENCES

1. Voigt P, von dem Bussche A. *The EU General Data Protection Regulation (GDPR): A Practical Guide.* Springer, Cham. 2017.
2. Government of India. *The Digital Personal Data Protection Act, 2023.* Ministry of Law and Justice, New Delhi. 2023.
3. McMahan B, Moore E, Ramage D, Hampson S, y Arcas BA. Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), PMLR.* 2017; 54: 1273–1282.
4. Kairouz P, McMahan HB, Avent B, et al. Advances and open problems in federated learning. *Foundations and Trends in Machine Learning.* 2021; 14(1–2): 1–210.
5. Yang Q, Liu Y, Chen T, Tong Y. Federated machine learning: concept and applications. *ACM Transactions on Intelligent Systems and Technology.* 2019; 10(2): Article 12.
6. Li T, Sahu AK, Talwalkar A, Smith V. Federated learning: challenges, methods, and future directions. *IEEE Signal Processing Magazine.* 2020; 37(3): 50–60.
7. Zhu L, Liu Z, Han S. Deep leakage from gradients. *Advances in Neural Information Processing Systems.* 2019; 32.
8. Geiping J, Bauermeister H, Dröge H, Moeller M. Inverting gradients – how easy is it to break privacy in federated learning? *Advances in Neural Information Processing Systems.* 2020; 33: 16937–16947.
9. Shokri R, Stronati M, Song C, Shmatikov V. Membership inference attacks against machine learning models. *IEEE Symposium on Security and Privacy.* 2017; 3–18.
10. Bagdasaryan E, Veit A, Hua Y, Estrin D, Shmatikov V. How to backdoor federated learning. *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics, PMLR.* 2020; 108: 2938–2948.
11. Dwork C, Roth A. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science.* 2014; 9(3–4): 211–407.
12. Abadi M, Chu A, Goodfellow I, McMahan HB, Mironov I, Talwar K, Zhang L. Deep learning with differential privacy. *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS).* 2016; 308–318.

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

13. Bonawitz K, Ivanov V, Kreuter B, et al. Practical secure aggregation for privacy-preserving machine learning. *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS).* 2017; 1175–1191.
14. Gentry C. Fully homomorphic encryption using ideal lattices. *Proceedings of the 41st ACM Symposium on Theory of Computing (STOC).* 2009; 169–178.
15. Rieke N, Hancox J, Li W, et al. The future of digital health with federated learning. *npj Digital Medicine.* 2020; 3: 119.
16. Hard A, Rao K, Mathews R, et al. Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604.* 2018.
17. Lim WYB, Luong NC, Hoang DT, et al. Federated learning in mobile edge networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials.* 2020; 22(3): 2031–2063.
18. Zhao Y, Li M, Lai L, Suda N, Civin D, Chandra V. Federated learning with non-IID data. *arXiv preprint arXiv:1806.00582.* 2018.
19. Hsu TMH, Qi H, Brown M. Measuring the effects of non-identical data distribution for federated visual classification. *arXiv preprint arXiv:1909.06335.* 2019.
20. Li T, Sahu AK, Zaheer M, Sanjabi M, Talwalkar A, Smith V. Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems.* 2020; 2: 429–450.
21. Karimireddy SP, Kale S, Mohri M, Reddi S, Stich S, Suresh AT. SCAFFOLD: Stochastic controlled averaging for federated learning. *Proceedings of the 37th International Conference on Machine Learning, PMLR.* 2020; 119: 5132–5143.
22. Fallah A, Mokhtari A, Ozdaglar A. Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach. *Advances in Neural Information Processing Systems.* 2020; 33: 3557–3568.
23. Konečný J, McMahan HB, Yu FX, Richtárik P, Suresh AT, Bacon D. Federated learning: strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492.* 2016.
24. Sattler F, Wiedemann S, Müller KR, Samek W. Robust and communication-efficient federated learning from non-i.i.d. data. *IEEE Transactions on Neural Networks and Learning Systems.* 2020; 31(9): 3400–3413.
25. Caldas S, Duddu SMK, Wu P, et al. LEAF: A benchmark for federated settings. *arXiv preprint arXiv:1812.01097.* 2018.

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

26. Mothukuri V, Parizi RM, Pouriyeh S, Huang Y, Dehghantanha A, Srivastava G. A survey on security and privacy of federated learning. *Future Generation Computer Systems.* 2021; 115: 619–640.
27. Lyu L, Yu H, Yang Q. Threats to federated learning: A survey. *arXiv preprint arXiv:2003.02133.* 2020.
28. Blanchard P, El Mhamdi EM, Guerraoui R, Stainer J. Machine learning with adversaries: Byzantine tolerant gradient descent. *Advances in Neural Information Processing Systems.* 2017; 30.
29. Mironov I. Rényi differential privacy. *Proceedings of the IEEE 30th Computer Security Foundations Symposium (CSF).* 2017; 263–275.
30. McMahan HB, Ramage D, Talwar K, Zhang L. Learning differentially private recurrent language models. *International Conference on Learning Representations (ICLR).* 2018.
31. Wei K, Li J, Ding M, et al. Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security.* 2020; 15: 3454–3469.
32. Mohassel P, Zhang Y. SecureML: A system for scalable privacy-preserving machine learning. *IEEE Symposium on Security and Privacy.* 2017; 19–38.
33. Li Q, Wen Z, Wu Z, et al. A survey on federated learning systems: Vision, hype and reality for data privacy and protection. *IEEE Transactions on Knowledge and Data Engineering.* 2023; 35(4): 3347–3366.
34. Kitchenham B, Charters S. *Guidelines for performing systematic literature reviews in software engineering (EBSE Technical Report EBSE-2007-01).* Keele University and Durham University. 2007.
35. Sheller MJ, Edwards B, Reina GA, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports.* 2020; 10: 12598.

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved

***Author for Correspondence:**
Sumit K. Sengupta
E-mail: sumit.k.sengupta45@gmail.com

[1]Assistant Professor, Department of Computer Science and Engineering, Ambalika Institute of Management & Technology, Lucknow
[2]Associate Professor, Department of Computer Science and Engineering, Ambalika Institute of Management & Technology, Lucknow
[3]Research Scholar, Department of Computer Science and Engineering, Ambalika Institute of Management & Technology, Lucknow

Received Date: May 17, 2025
Accepted Date: May 30, 2025
Published Date: June 14, 2025

Citation: Sumit K. Sengupta, Moumita D. Pal, Subhajit R. Dutta. Federated Learning and Privacy-Preserving AI: Emerging Approaches for Distributed Intelligent Computing — A Systematic Review of Algorithms, Privacy Mechanisms, and Deployment Trade-offs. Journal of Converging Technologies in Intelligent. 2025; 1(2): 118-132p.

Page 118-132 © MANTECH PUBLICATIONS 2025. All Rights Reserved