Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

Review Article

A Systematic Review of Federated Learning: Emerging Techniques, Challenges, and Research Directions

[1]Mohammed Abdul Ameer Jabbar
Department of Scientific Affairs
University of Information Technology and Communications
Baghdad, Iraq
mohammedaji@uoitc.edu.iq

[2]Ahmed Sami Jaddoa
Department of Informatics Systems Management
University of Information Technology and Communications
Baghdad, Iraq
ahmed.sami@uoitc.edu.iq

[3]Usama Samir Mahmoud
University of Information Technology and Communications
Baghdad, Iraq
usama.s.mahmoud@uoitc.edu.iq

ARTICLE INFO
Article History
Received: 27/07/2025
Accepted: 31/08/2025
Published: 07/09/2025

ABSTRACT

Federated Learning FL is a rapidly evolving machine learning paradigm that enables collaborative model training across decentralized data sources while preserving data privacy. Since its inception in 2016, FL has emerged as a transformative approach in domains such as healthcare, IoT, and edge computing, where data sensitivity and regulatory constraints limit centralized processing. This systematic review consolidates findings from 50 high-quality studies selected from over 250 papers to present a comprehensive synthesis of FL methodologies, core aggregation techniques, privacy-preserving mechanisms, security threats, domain-specific applications, and emerging trends. We categorize challenges into communication overhead, heterogeneity (device, data, model), fairness, trust, and evaluation inconsistencies. We highlight research gaps, notably in standardized evaluation, incentive mechanisms, and deployment scalability. By reviewing recent advances such as vertical federated learning, federated unlearning, and blockchain-based incentives, this paper offers a roadmap for future research and identifies open questions vital for widespread FL adoption.

Keywords: *Federated Learning; Decentralized And Distributed FL; Federated Learning Aggregation Methods; Data Privacy And Client Selection Methods; Federated Learning Applications*

1. INTRODUCTION

The increase volume of data generated across distributed sources such as smartphones, wearable devices, sensors and hospital systems has catalyzed the demand for privacy-preserving, decentralized Machine Learning ML solutions. Traditional centralized ML pipelines require transferring all data to central server which introduces risks related to privacy violations, regulatory non-compliance for instance (GDPR, HIPAA) and communication bottlenecks. In response, FL was proposed by Google in 2016 as a decentralized alternative that allows multiple clients (devices, institutions) to collaboratively train a shared model while keeping data localized [1-3]. The motivation for FL is multifaceted. First and foremost, FL addresses the critical need for data privacy and security in sectors where data sensitivity is paramount. By keeping data on local devices or within institutional boundaries, FL minimizes the risk of data breaches and unauthorized access, thereby supporting compliance with regulations such as the General Data Protection Regulation GDPR in Europe and the Health Insurance Portability and Accountability Act HIPAA in the United States. In addition to privacy, FL also enable organizations to leverage the collective intelligence embedded in distributed dataset, which can lead to more robust and generalizable model. This is particularly important in healthcare, where the ability to train models on diverse patient population can improve the accuracy and fairness of predictive analytics, as demonstrated in recent studies on prognosis assessment for acute pulmonary thromboembolism. Another key motivation is the practical challenge of data movement and storage. As data volumes continue to grow, the costs and logistical complexities associated with transferring large dataset to central location become prohibitive. FL offers scalable solution by reducing the need for data movement instead relying on the exchange of lightweight model updates. FL operates by enabling local training on client devices and only sharing model updates (e.g., gradients or parameters) with a central or peer-coordinated aggregator. This paradigm shift supports privacy-aware learning and has gained traction in sensitive domains like healthcare [4, 5], financial systems, and autonomous vehicles. Despite its

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

transformative potential, FL introduces new technical and practical challenges such as client and data heterogeneity, high communication costs, vulnerability to adversarial attacks, fairness, scalability and evaluation difficulties [6]. To address these challenges, researchers have proposed several methods, from core aggregation algorithms like FedAvg and secure aggregation to client selection policies [7] and decentralized trust mechanisms like blockchain. Concurrently, emerging FL extensions such as Vertical Federated Learning VFL, federated unlearning, and incentive models are reshaping the research landscape [8].

The aim of this review is to provide comprehensive and systematic synthesis of the current state of FL, with a particular focus on its foundational methods, the challenges it faces and the future direction that are shaping ongoing research and deployment. This review seeks to elucidate the core of FL methods and architectures, such as federated averaging, secure aggregation, and distributed client selection optimization. In doing so, the review not only maps the technical landscape of FL, but also contextualizes its evolution within broader societal and regulatory trends. Our contributions are summarized as follows:

- Synthesis of core FL methods and aggregation strategies.
- Detailed classification of FL challenges and security issues.
- Survey of domain-specific applications in healthcare, IoT, and edge computing.
- Coverage of emerging research trends such as federated unlearning and incentive frameworks.
- Evaluation of current benchmarking practices and research gaps.
- Roadmap of open research questions.

The rest of the paper is organized as Section 2 reviews recent FL literatures, Section 3 includes the core principles and aspects in FL, Section 4 describes the FL architecture, Section 5 involves client selection and optimization methods, section 6 clarifies discussions, and Section 7 illustrates conclusions.

2. RECENT FEDERATED LEARNING LITERATURE STUDIES

This literature survey investigates the recent advancements in FL by analyzing 19 foundational and influential review papers from 2019 to 2024. The surveyed literature spans diverse aspects of FL. It is presented as ranging from general reviews and lifecycle overviews to focused studies on specific challenges such as security, privacy, and heterogeneity. Notably, works by Lo [9], Zhang [10] and Liu [4] provide comprehensive surveys and taxonomies outlining the progression of FL across architectures and aggregation strategies. Healthcare is dominant area of focus, as seen in studies by Liu [3], Sheller [11], and Zhang [12], which emphasize FL potential for privacy preserving collaboration across medical institutions. Likewise, the significance of federated learning in the Internet of Things and edge computing is examined in studies such as Khan [13] and Shaheen [14]. Security is a fundamental aspect of FL research, as emphasized by Rodriguez-Barroso [15] and Mothukuri [16], who classify threat models and defenses. Furthermore, recent initiatives focus on heterogeneity and robustness, as examined by Ye [17] and Huang [18], highlighting the necessity for adaptable, generalizable models. Table 1 illustrates the recent federated learning literature studies and insights, and Figure 1 shows the distribution of FL studies across various domains.

TABLE I. Recent FL literature studies and insights

Year of Publication	Cover Fields	Area of Study	Ref.
2019	Broad overview of FL challenges, methods, and future research directions.	General / Core FL	[19]
2020	Reviews 231 studies, covering FL lifecycle, challenges.	General FL	[9]
2020	Reviews FL for IoT, including privacy and scalability.	Internet of Things	[13]
2020	Demonstrates FL's effectiveness in multi-institutional medical collaborations.	Healthcare	[11]
2020	Shows FL can match centralized models in health research while preserving privacy.	Healthcare	[20]
2020	Reviews FL research activities, applications, and taxonomy.	General FL	[21]

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

Year of Publication	Cover Fields	Area of Study	Ref.
2021	Systematic survey on FL, including data partitioning, privacy, models, and challenges.	General FL	[10]
2021	Comprehensive review of FL security and privacy issues and future research.	Security & Privacy	[16]
2022	Surveys FL models for cyber_security, privacy, and attack detection.	Cybersecurity	[22]
2022	Reviews FL applications in edge networks, IoT, healthcare, and more.	Multiple Domains (IoT, Healthcare)	[14]
2022	Reviews FL threats, attacks, defenses, and experimental findings.	Security & Threats	[15]
2023	Focuses on heterogeneity in FL and solutions at data, model, and server levels.	Heterogeneous FL	[17]
2023	Systematic review of recent FL methods, applications, and frameworks.	General FL	[3]
2023	Reviews FL advances in generalization, robustness, and fairness, with benchmarks.	Model Evaluation / Fairness	[18]
2023	Reviews FL methods and applications in medical image analysis.	Healthcare (Medical Imaging)	[4]
2023	Reviews FL frameworks, applications, and privacy concerns.	General FL	[23]
2023	Reviews FL challenges, aggregation techniques, and development tools.	Development Tools / Aggregation	[2]
2024	Analyzes FL in healthcare, highlighting methodological flaws and recommendations.	Healthcare	[24]
2024	Reviews new FL methodologies for healthcare and systemic issues.	Healthcare	[12]

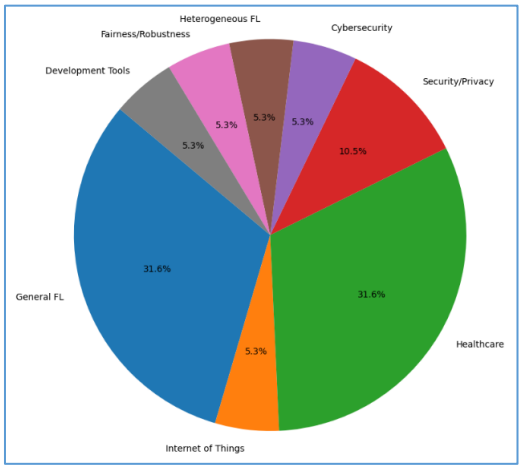

Fig. 1. Distribution of the FL study percentage domains

This figure shows the FL studies and distributions across various domains. It presents the data as percentages obtained from recent studies in the literature that describe the number of published studies in that field. The analyzed studies collectively highlight the dynamic progression of FL, its increasing applications, and the urgent challenges of

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

implementation in practical environments. Despite persistent challenges in communication expenses, system diversity, and equity, the literature indicates a dynamic and promising research direction, especially in critical areas such as healthcare, IoT and cyber security.

3. FEDERATED LEARNING CORE PRINCIPLES AND ASPECTS

The FL is built on several core principles that distinguish it from traditional centralized ML and enable privacy-preserving as well as collaborative model development across multiple institutions or devices

3.1 Data Privacy and Local Training

The raw data in FL never leaves the local device or institution. However, each participant trains the model on its own data and only share updates from the model which includes (weights or gradients), with central server or aggregator. This approach protects sensitive information and complies with privacy regulations, making FL especially suitable for domains like healthcare [25].

3.2 Collaborative Model Building, Aggregation and Iterative Methods

Aggregation is a critical aspect in FL model updates. The central server collects model updates from all participants and aggregates them to form a global model. This process is repeated over several rounds, allowing the global model to benefit from the collective knowledge of all participants without exposing individual data [26]. The iterative methods improvement encourages the global model to be redistributed to participants for further local training, and the cycle continues, improving the model's performance with each round [27].

3.3 Handling Data Heterogeneity, Security Threats

For supporting diverse data distributions, FL is designed to work with data that may be Non-Independent and Identically Distributed (Non-IID) across clients. This means each participant's data can have different characteristics, and the FL process still aims to build a robust, generalizable model [28]. Figure 2 shows the aggregation and security issues on FL distributed systems.

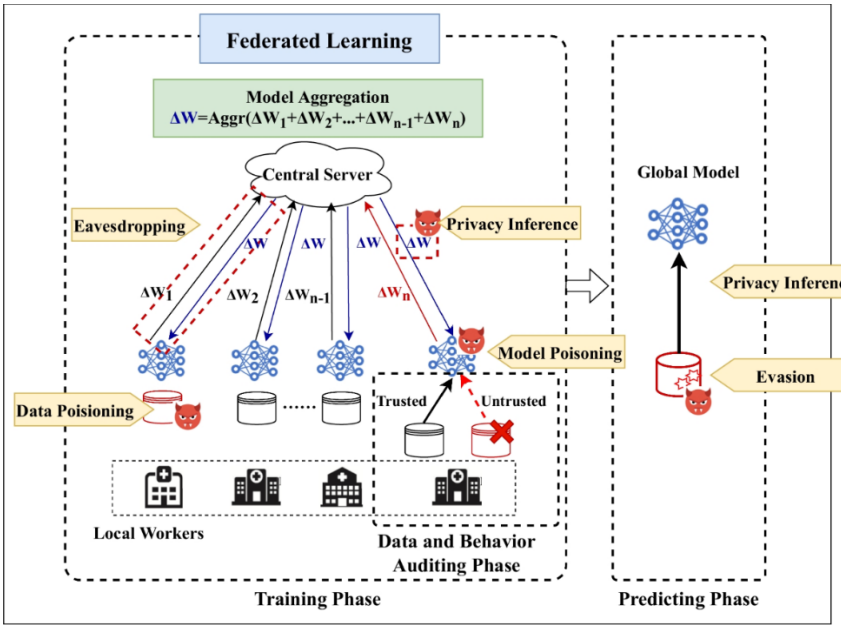

Fig. 2. Federated learning aggregations and security threads [29]

Securing collaboration is ensured in FL model. This model shares only the model parameters and not raw data, FL reduces the risk of data breaches. Additional security measures, such as secure aggregation protocols, can further protect the integrity and confidentiality of model updates.

4. FEDERATED LEARNING ARCHITECTURE

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

Federated Learning FL enables collaborative training of machine learning models among various decentralized clients, while safeguarding data privacy and adhering to regulatory standards. In contrast to conventional centralized methods that aggregate raw data in a central server, federated learning allows data proprietors to maintain local datasets while only disseminating model parameters. This is especially advantageous in privacy-sensitive sectors such as healthcare, finance, and smart infrastructure, where data is regulated by stringent policies like HIPAA and GDPR [30, 31]. The fundamental architecture of FL consists of multiple interconnected components that guarantee secure, scalable, and efficient learning. At the periphery of the system are the local client devices or entities such as hospitals, smartphones, or sensors that possess proprietary or sensitive datasets [32]. These clients conduct model training locally and calculate updates, such as gradients or model weights, without disclosing raw data to any central authority. In a healthcare context, hospitals train localized models using patient data and transmit solely the encrypted model updates to a centralized server, thereby substantially mitigating the risk of data leakage. Figure 3 shows the FL Distributed training and Data privacy.

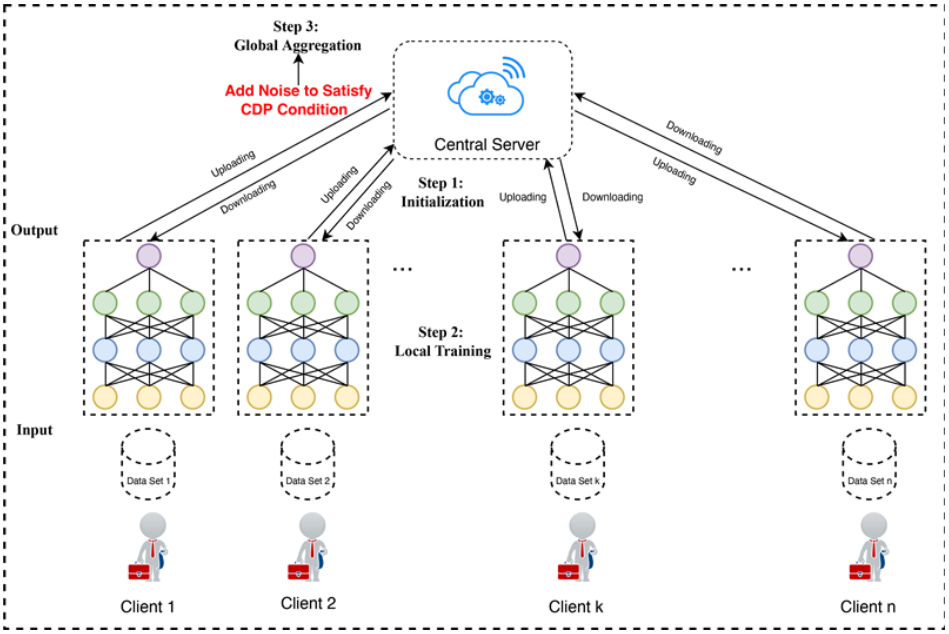

Fig. 3. Distributed training and Data privacy of Federated learning [33]

The central server or aggregator is the focal point of the system. This server gathers model updates from various clients and consolidates them, typically employing the Federated Averaging (FedAvg) algorithm [34], to generate a new global model. The global model is subsequently redistributed to the clients for the subsequent training round. This iterative communication facilitates ongoing model enhancement while maintaining data locality. Advanced implementations of the central aggregator frequently incorporate secure aggregation protocols to prevent the reverse-engineering of individual model updates, thereby safeguarding client data [35].

A vital element is the communication layer, which facilitates the secure and efficient transmission of model updates. Due to bandwidth constraints and the diversity of edge environments, communication protocols need optimization to facilitate asynchronous updates, tolerate latency, and implement compression methods like sparsification or quantization. This is particularly crucial in mobile or IoT-based federated learning applications [36], where clients may experience intermittent connectivity or operate under constraints of limited power and network access. The model update and aggregation mechanisms delineate the process by which client updates are amalgamated to create the global model. Although FedAvg is the predominant method, alternative algorithms like FedProx, FedNova [37], and Scaffold have been formulated to tackle challenges associated with data heterogeneity, system variability, and convergence instability. These methodologies seek to enhance model robustness in contexts as a common occurrence in practical applications where client datasets are Non-Independent and Identically Distributed (Non-IID). For instance, federated learning has been utilized in personalized healthcare, where patient demographics and disease patterns vary among institutions, rendering centralized learning suboptimal or potentially biased [38].

Furthermore, federated learning frameworks can be augmented with advancements like differential privacy, Secure MultiParty Computation SMPC, and blockchain integration to enhance security, transparency, and trust. Differential privacy incorporates calibrated noise into model updates to thwart re-identification attacks, whereas SMPC facilitates

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

computations on encrypted data among multiple parties. Conversely, blockchain offers an immutable ledger of all model updates, facilitating the verifiability and auditability of contributions an area of increasing significance in critical sectors such as finance and healthcare [39, 40]. Table 2 shows a summary of the core component FL architecture.

TABLE II. Summary of the Core Components of Federated Learning Architecture

Component	Function	Example in Practice	Relevant Methods
Local Clients	Trains models on private local data and generates updates	Hospitals, smartphones, wearable devices	Local SGD, Personalization
Central Aggregator	Collects and merges model updates into a global model	Cloud server, institutional aggregator	FedAvg, FedProx, FedNova
Communication Layer	Secures transmission of model updates; handles delays and compression	Encrypted TLS channels, 5G, WiFi networks	Update Compression, DP, SMPC
Aggregation Mechanism	Combines model updates, adapts to data/system heterogeneity	Statistical weighting, secure protocols	FedAvg, Secure Aggregation
Privacy Enhancements	Ensures confidentiality during training and aggregation	GDPR/HIPAA-compliant federated systems	Differential Privacy, SMPC

The architecture of federated learning consists of a complex yet efficient collaboration among edge clients, a coordinating server, secure communication channels, and robust aggregation algorithms. It provides a robust solution to the urgent demand for privacy-preserving collaborative AI in decentralized and sensitive contexts. FL consistently showcases its utility and versatility in both research and practical applications, as demonstrated by its successful use in prognosticating acute pulmonary thromboembolism [41] and predicting the progression of diabetic kidney disease [42, 43].

5. FEDERATED LEARNING CLIENT SELECTION METHODS

Client selection methods are a core component of the FL, directly affecting model accuracy, training efficiency, and system scalability. Because clients (devices or nodes) differ in data quality, computational power, and network reliability, choosing the right subset of clients for each training round is essential for optimal FL performance. The categories of client selection methods are including (random selection, resource aware selection [44], data/contributions-based selection, reputation and reliability-based selection and hybrid and adaptive methods [45]). In random selection, clients are chosen randomly for each round. This is simple but can lead to poor performance due to the heterogeneity of the device and data. In resource-aware selection, methods look at the capabilities of the devices (CPU, memory, battery) and the state of the network to choose clients that can train and communicate quickly. This cuts down on stragglers and speeds things up overall [46-49]. In selection based on data or contributions, clients are chosen based on how good, plentiful, or varied their local data is, or how much they have helped to improve the model in the past. The goal of this method is to make the model as accurate and generalizable as possible. In reputation and reliability-based selection, clients who have a history of reliable participation or high-quality updates are given priority. This makes the system more robust and trustworthy. The hybrid and adaptive methods, which combine several criteria (like resource, data, and reputation) and change to fit new situations, find a balance between efficiency, fairness, and accuracy [50-52]. Table 3 shows summary of the main client selection methods in FL.

TABLE III. Summary of the main client selection methods

Method Type	Main Focus	Key Benefits	Ref.
Random	Simplicity	Easy to implement	[53]
Resource-Aware	Device/network status	Reduces stragglers, faster	[46-49]
Data-Based	Data quality/diversity	Improves model accuracy	[46-49]
Reputation-Based	Reliability/trust	Robustness, security	[49]
Hybrid/Adaptive	Multiple criteria	Balances goals, flexible	[51]

5.1 Challenges in Client Selection Methods

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

There are many ways to choose clients in federated learning, from simple random methods to more complex adaptive strategies that take into account the device's resources, the data value and its reliability. The method you choose has a big effect on FL efficiency, fairness, and model quality. Researchers are still working to improve these methods for use in a variety of real-world situations.

5.2 Client Selection Optimization

Optimizing client selection is a critical aspect of federated learning, significantly influencing model efficacy, training efficiency, and communication expenses. Due to the variability in data quality, computational resources, and network conditions among clients (devices or nodes), it is crucial to select the appropriate subset of clients for each training round to ensure the effectiveness and efficiency of FL [49, 54]. The primary aims of client selection encompass:

- Optimizing model efficacy, and choosing clients with varied and superior data can enhance the precision and applicability of the global model.
- Minimizing communication and computational expenses, and selecting a limited group of suitable clients diminishes bandwidth consumption and computational demands, thereby enhancing the scalability and feasibility of machine learning.
- Managing variability, and optimized selection addresses discrepancies in client resources, data distributions, and availability, which are prevalent variations in real-world federated learning scenarios.

5.3 Do Client Selection Strategies Improve Federated Learning Efficiency?

Client selection methods are significantly improving the efficiency of FL by addressing challenges that related to device heterogeneity, resource constraint, data quality and communication overhead. Rather than randomly selecting clients, optimized strategies ensure that the most suitable clients participate in each training round, leading to faster convergence, reduced resource waste and better model performance. Random client selection can harm efficiency and model quality due to device and data heterogeneity. Strategic selection methods (resource-aware, data-driven, reputation-based, or hybrid) consistently outperform random selection in both speed and accuracy. Trade-offs exist between maximizing accuracy, minimizing training time, and ensuring fairness, but optimizing selection [50]. Optimized client selection strategies are essential for efficient FL, which reduces training time, lowers communication costs, and improves model quality by ensuring that the most appropriate clients participate in each round, making FL more practical and scalable in real-world applications.

6. DISCUSSIONS

The literature indicates that FL methods have significantly advanced, with robust algorithms and supporting infrastructure now facilitating real-world implementations across various institutions and countries. For instance, FL-based models have been effectively utilized for prognostic evaluation in acute pulmonary thromboembolism, demonstrating robust predictive capabilities and surpassing conventional centralized models across various assessment metrics. These practical applications demonstrate FL's capacity to enable multi-institutional collaboration while safeguarding patient privacy and data security.

FL has made significant advancements in recent years. However, it still faces enduring and complex challenges that limit its scalability and generalizability. A major concern is the heterogeneity of data and systems. In practice, data gathered by various institutions or devices frequently exhibits discrepancies in distribution, quality, and format (Non-IID data), which can adversely affect model convergence and overall performance. The heterogeneity of systems, including variations in computational resources and network connectivity among clients, exacerbates the challenges of coordination and efficiency in the federated learning process. Communication overhead constitutes a significant obstacle, as the continual transmission of model updates between clients and the central server can burden bandwidth and elevate latency, especially in large-scale or resource-limited settings.

Although reduction is relative to centralized methods, privacy risks in federated learning are not completely eradicated. There exists a risk of information leakage during model updates, and adversarial attacks aimed at the aggregation process or exploiting weaknesses in the communication protocol continue to be significant concerns. These challenges highlight the necessity for ongoing research into sophisticated privacy-preserving methodologies, including secure multiparty computation and differential privacy, to enhance the security assurances of federated learning systems.

The evidence substantiating the feasibility of FL and its capacity to attain performance on par with, or surpassing, as centralized models are robust, especially in healthcare applications where direct comparisons have been conducted. Nevertheless, the literature underscores the imperative for additional investigation in several critical domains. There

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

is an urgent necessity to devise and thoroughly assess sophisticated privacy methodologies that can be effortlessly incorporated into federated learning workflows. Moreover, concerns regarding equity ensuring that FL models function impartially across varied populations and institutions necessitate increased focus. Ultimately, broadening the application of FL to encompass a wider array of real-world contexts, beyond its present emphasis on healthcare, will be crucial for unlocking its complete potential and tackling the distinct challenges posed by various sectors.

6.1 Research Gaps

Notwithstanding considerable advancements, research gaps persist in the implementation of sophisticated privacy-preserving methodologies, addressing extreme heterogeneity, and executing federated learning in varied, real-world contexts beyond healthcare and finance. Figure 4 shows matrix of research topics and study attributes showing coverage and gaps in FL.

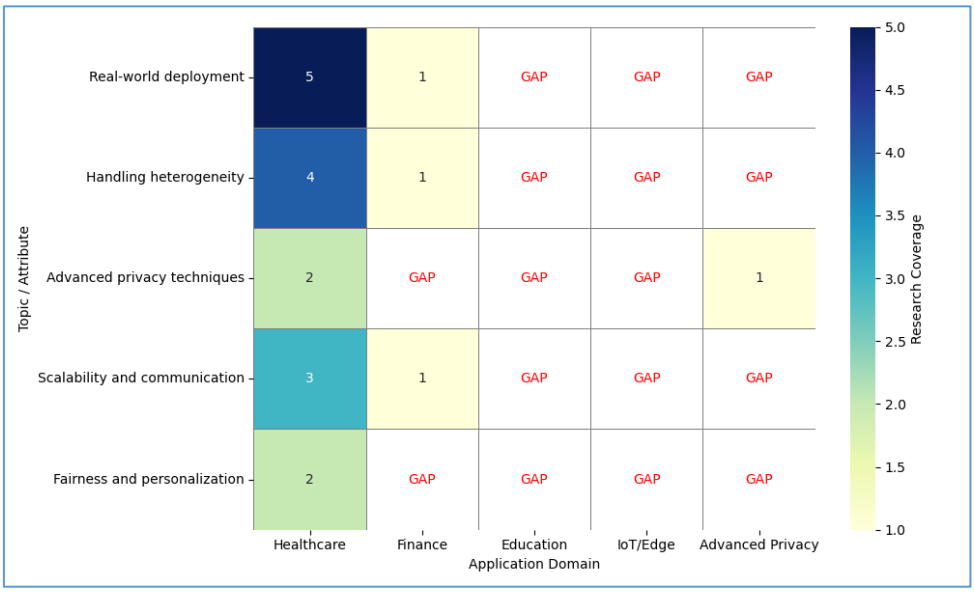

Fig. 4. Matrix of research topics and study attributes showing coverage and gaps in FL literature.

6.2 Open Research Questions

Future research should address the following open questions to advance FL's scalability, privacy, and applicability:

- How can federated learning be made robust to extreme statistical and system heterogeneity in real-world deployments? Addressing heterogeneity is crucial for FL's scalability and generalizability across diverse clients and environments.
- What are the most effective ways to integrate advanced privacy-preserving techniques into practical FL systems? Stronger privacy guarantees are needed to meet regulatory requirements and build trust in FL applications.
- How can FL be efficiently deployed in domains beyond healthcare and finance, such as IoT, education, and edge computing? Expanding FL's application scope will maximize its societal impact and reveal new technical challenges.

7. CONCLUSIONS

The FL exhibited significant potential in fields such as healthcare, IoT and finance, facilitating collaborative model development while keeping data privacy. However, numerous enduring challenges such as communication, Non-IID data distribution, equity in client selection and susceptibility to adversarial threats continue to impede widespread adoption. This review emphasizes deficiencies in standardized evaluation metrics, scalability of deployment, and the creation of incentive mechanisms for ongoing client engagement. The domain is progressing with innovative developments including vertical federated learning, federated unlearning, blockchain-integrated trust systems, and adaptive client selection frameworks. These innovations indicate that federated learning is evolving from a theoretical

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

concept to a practical framework for secure, distributed artificial intelligence. Ensuring equity, trust, and efficiency in varied real-world contexts necessitates ongoing research into effective privacy-preserving methods, dependable client orchestration, and scalable system architecture. This review consolidates the current state of federated learning research and offers a prospective roadmap for addressing unresolved enquiries. The maturation of FL will hinge on its adherence to ethical AI principles, regulatory compliance and technical robustness to fully harness its potential in data-rich to the global ecosystems, and FL transition from promising prototypes to large-scale, trusted deployments can be only by addressing these challenges. As such, its future success will not only redefine how data driven intelligence is shared, but also shape the foundations of a more secure and collaborative digital society.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

No research funding for this paper.

References

[1] S. Banabilah, M. Aloqaily, E. Alsayed, N. Malik, and Y. Jararweh, "Federated learning review: Fundamentals, enabling technologies, and future applications," *Inf. Process. Manag.,* vol. 59, p. 103061, 2022-11-01 2022, doi: 10.1016/j.ipm.2022.103061.

[2] B. S. Guendouzi, S. Ouchani, H. E. Assaad, and M. E. Zaher, "A systematic review of federated learning: Challenges, aggregation methods, and development tools," *J. Netw. Comput. Appl.,* vol. 220, p. 103714, 2023-08-01 2023, doi: 10.1016/j.jnca.2023.103714.

[3] B. Liu, N. Lv, Y. Guo, and Y. Li, "Recent Advances on Federated Learning: A Systematic Survey," *Neurocomputing,* vol. 597, p. 128019, 2023-01-03 2023, doi: 10.48550/arXiv.2301.01299.

[4] M. Liu, "Federated Learning for Medical Image Analysis: A Survey," *Pattern recognition,* vol. 151, 2023-06-09 2023, doi: 10.48550/arXiv.2306.05980.

[5] R. S. Antunes, C. A. Da Costa, A. Küderle, I. A. Yari, and B. Eskofier, "Federated Learning for Healthcare: Systematic Review and Architecture Proposal," *ACM Transactions on Intelligent Systems and Technology (TIST),* vol. 13, pp. 1-23, 2022-02-04 2022, doi: 10.1145/3501813.

[6] A. Tariq *et al.*, "Trustworthy Federated Learning: A Comprehensive Review, Architecture, Key Challenges, and Future Research Prospects," *IEEE Open Journal of the Communications Society,* vol. 5, pp. 4920-4998, 2024-01-01 2024, doi: 10.1109/ojcoms.2024.3438264.

[7] N. Romandini, A. Mora, C. Mazzocca, R. Montanari, and P. Bellavista, "Federated Unlearning: A Survey on Methods, Design Guidelines, and Evaluation Metrics," *IEEE transactions on neural networks and learning systems,* vol. PP, 2024-01-10 2024, doi: 10.1109/TNNLS.2024.3478334.

[8] M. Ye, W. Shen, B. Du, E. Snezhko, V. Kovalev, and P. Yuen, "Vertical Federated Learning for Effectiveness, Security, Applicability: A Survey," *ACM Computing Surveys,* 2024-05-25 2024, doi: 10.48550/arXiv.2405.17495.

[9] S. K. Lo, Q. Lu, C. Wang, H.-Y. Paik, and L. Zhu, "A Systematic Literature Review on Federated Machine Learning," *ACM Computing Surveys (CSUR),* vol. 54, pp. 1-39, 2020-07-22 2020, doi: 10.1145/3450288.

[10] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, "A survey on federated learning," *Knowl. Based Syst.,* vol. 216, p. 106775, 2021-01-21 2021, doi: 10.1016/j.knosys.2021.106775.

[11] M. Sheller *et al.*, "Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data," *Scientific Reports,* vol. 10, 2020-07-28 2020, doi: 10.1038/s41598-020-69250-1.

[12] F. Zhang *et al.*, "Recent methodological advances in federated learning for healthcare," *Patterns,* vol. 5, 2024-06-01 2024, doi: 10.1016/j.patter.2024.101006.

[13] L. Khan, W. Saad, Z. Han, E. Hossain, and C. Hong, "Federated Learning for Internet of Things: Recent Advances, Taxonomy, and Open Challenges," *IEEE Communications Surveys & Tutorials,* vol. 23, pp. 1759-1799, 2020-09-28 2020, doi: 10.1109/COMST.2021.3090430.

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

[14] M. Shaheen, M. Farooq, T. Umer, and B.-S. Kim, "Applications of Federated Learning; Taxonomy, Challenges, and Research Trends," *Electronics,* 2022-02-21 2022, doi: 10.3390/electronics11040670.

[15] N. Rodriguez-Barroso, D. J. Lopez, M. Luzon, F. Herrera, and E. Martínez-Cámara, "Survey on Federated Learning Threats: concepts, taxonomy on attacks and defences, experimental study and challenges," *ArXiv,* vol. abs/2201.08135, 2022-01-20 2022, doi: 10.1016/j.inffus.2022.09.011.

[16] V. Mothukuri, R. Parizi, S. Pouriyeh, Y.-P. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Gener. Comput. Syst.,* vol. 115, pp. 619-640, 2021-02-01 2021, doi: 10.1016/j.future.2020.10.007.

[17] M. Ye, X. Fang, B. Du, P. Yuen, and D. Tao, "Heterogeneous Federated Learning: State-of-the-art and Research Challenges," *ACM Computing Surveys,* vol. 56, pp. 1-44, 2023-07-20 2023, doi: 10.1145/3625558.

[18] W. Huang *et al.*, "Federated Learning for Generalization, Robustness, Fairness: A Survey and Benchmark," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 46, pp. 9387-9406, 2023-11-12 2023, doi: 10.1109/TPAMI.2024.3418862.

[19] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated Learning: Challenges, Methods, and Future Directions," *IEEE Signal Processing Magazine,* vol. 37, pp. 50-60, 2019-08-21 2019, doi: 10.1109/MSP.2020.2975749.

[20] J. Hernandez *et al.*, "Privacy-first health research with federated learning," *NPJ Digital Medicine,* vol. 4, 2020-12-24 2020, doi: 10.1038/s41746-021-00489-2.

[21] L. Li, Y. Fan, and K.-Y. Lin, "A Survey on federated learning*," *2020 IEEE 16th International Conference on Control & Automation (ICCA),* pp. 791-796, 2020-10-09 2020, doi: 10.1109/ICCA51439.2020.9264412.

[22] M. Alazab, S. Rm, P. M, P. K. R. Maddikunta, T. Gadekallu, and V. Q. Pham, "Federated Learning for Cybersecurity: Concepts, Challenges, and Future Directions," *IEEE Transactions on Industrial Informatics,* vol. 18, pp. 3501-3509, 2022-05-01 2022, doi: 10.1109/tii.2021.3119038.

[23] H. Kaur, V. Rani, M. Kumar, M. Sachdeva, A. Mittal, and K. Kumar, "Federated learning: a comprehensive review of recent advances and applications," *Multim. Tools Appl.,* vol. 83, pp. 54165-54188, 2023-11-30 2023, doi: 10.1007/s11042-023-17737-0.

[24] M. Li, P. Xu, J. Hu, Z. Tang, and G. Yang, "From Challenges and Pitfalls to Recommendations and Opportunities: Implementing Federated Learning in Healthcare," *Medical image analysis,* vol. 101, p. 103497, 2024-09-15 2024, doi: 10.48550/arXiv.2409.09727.

[25] R. Dembani, I. Karvelas, N. A. Akbar, S. Rizou, D. Tegolo, and S. Fountas, "Agricultural data privacy and federated learning: A review of challenges and opportunities," *Computers and Electronics in Agriculture,* vol. 232, p. 110048, 2025/05/01/ 2025, doi: https://doi.org/10.1016/j.compag.2025.110048.

[26] Y. Deng *et al.*, "FAIR: Quality-Aware Federated Learning with Precise User Incentive and Model Aggregation," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, 10-13 May 2021 2021, pp. 1-10, doi: 10.1109/INFOCOM42981.2021.9488743.

[27] K. Pillutla, S. M. Kakade, and Z. Harchaoui, "Robust Aggregation for Federated Learning," *IEEE Transactions on Signal Processing,* vol. 70, pp. 1142-1154, 2022, doi: 10.1109/TSP.2022.3153135.

[28] H. S. Sikandar, H. Waheed, S. Tahir, S. U. R. Malik, and W. Rafique, "A Detailed Survey on Federated Learning Attacks and Defenses," *Electronics,* vol. 12, no. 2, p. 260, 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/2/260.

[29] P. Liu, X. Xu, and W. Wang, "Threats, attacks and defenses to federated learning: issues, taxonomy and perspectives," *Cybersecurity,* vol. 5, no. 1, p. 4, 2022/02/02 2022, doi: 10.1186/s42400-021-00105-6.

[30] J. Zhou, X. Wang, Y. Li, Y. Yang, and J. Shi, "Federated-learning-based prognosis assessment model for acute pulmonary thromboembolism," *BMC Medical Informatics and Decision Making,* vol. 24, 2024-05-27 2024, doi: 10.1186/s12911-024-02543-x.

[31] X. Ouyang, "Design and Deployment of Multi-Modal Federated Learning Systems for Alzheimer's Disease Monitoring," *Proceedings of the 21st Annual International Conference on Mobile Systems, Applications and Services,* 2023-06-18 2023, doi: 10.1145/3581791.3597505.

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

[32] D. He *et al.*, "The use of artificial intelligence in the treatment of rare diseases: A scoping review," *Intractable & Rare Diseases Research,* vol. 13, no. 1, pp. 12-22, 2024, doi: 10.5582/irdr.2023.01111.

[33] Y. Zhang, Lu, Y. and Liu, F., "A Systematic Survey for Differential Privacy Techniques in Federated Learning," *Journal of Information Security,* vol. 14, pp. 111-135, 2023, doi: 10.4236/jis.2023.142008.

[34] T. Zhou, Z. Lin, J. Zhang, and D. H. K. Tsang, "Understanding and Improving Model Averaging in Federated Learning on Heterogeneous Data," *IEEE Transactions on Mobile Computing,* vol. 23, no. 12, pp. 12131-12145, 2024, doi: 10.1109/TMC.2024.3406554.

[35] Y. Suzuki *et al.*, "Heterogeneity analysis of acute exacerbations of chronic obstructive pulmonary disease and a deep learning framework with weak supervision and privacy protection," *bioRxiv,* p. 2023.12.04.570028, 2023, doi: 10.1101/2023.12.04.570028.

[36] N. Khajehali, J. Yan, Y.-W. Chow, and M. Fahmideh, "A Comprehensive Overview of IoT-Based Federated Learning: Focusing on Client Selection Methods," *Sensors,* vol. 23, no. 16, p. 7235, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/16/7235.

[37] M. Haipeng *et al.*, "Using federated learning technology to improve smart grid fault diagnosis efficiency and privacy protection," in *Proc.SPIE*, 2024, vol. 13269, p. 132690O, doi: 10.1117/12.3045669. [Online]. Available: https://doi.org/10.1117/12.3045669

[38] D. C. Nguyen *et al.*, "Federated learning for smart healthcare: A survey," *ACM Computing Surveys (Csur),* vol. 55, no. 3, pp. 1-37, 2022.

[39] J. Frizzo-Barker, P. Chow-White, P. Adams, J. Mentanko, D. Ha, and S. Green, "Blockchain as a disruptive technology for business: A systematic review," *Int. J. Inf. Manag.,* vol. 51, p. 102029, 2020-04-01 2020, doi: 10.1016/j.ijinfomgt.2019.10.014.

[40] F. Martin, T. Sun, and C. Westine, "A systematic review of research on online teaching and learning from 2009 to 2018," *Computers & Education,* vol. 159, pp. 104009-104009, 2020-09-09 2020, doi: 10.1016/j.compedu.2020.104009.

[41] J. Zhou, X. Wang, Y. Li, Y. Yang, and J. Shi, "Federated-learning-based prognosis assessment model for acute pulmonary thromboembolism," *BMC Medical Informatics and Decision Making,* vol. 24, no. 1, p. 141, 2024/05/27 2024, doi: 10.1186/s12911-024-02543-x.

[42] M. A. Mohammed *et al.*, "Federated-reinforcement learning-assisted IoT consumers system for kidney disease images," *IEEE Transactions on Consumer Electronics,* vol. 70, no. 4, pp. 7163-7173, 2024, doi: 10.1109/TCE.2024.3384455.

[43] J. M. Nandhini, S. Joshi, and K. Anuratha, "Federated Learning Based Prediction of Chronic Kidney Diseases," in *2022 1st International Conference on Computational Science and Technology (ICCST)*, 9-10 Nov. 2022 2022, pp. 1-6, doi: 10.1109/ICCST55948.2022.10040317.

[44] J. Wen, J. Zhang, Z. Zhang, Z. Cui, X. Cai, and J. Chen, "Resource-aware multi-criteria vehicle participation for federated learning in Internet of vehicles," *Information Sciences,* vol. 664, p. 120344, 2024/04/01/ 2024, doi: https://doi.org/10.1016/j.ins.2024.120344.

[45] M. Hamood, A. Albaseer, M. Abdallah, A. Al-Fuqaha, and A. Mohamed, "Optimized Federated Multitask Learning in Mobile Edge Networks: A Hybrid Client Selection and Model Aggregation Approach," *IEEE Transactions on Vehicular Technology,* vol. 73, no. 11, pp. 17613-17629, 2024, doi: 10.1109/TVT.2024.3427349.

[46] N. Khajehali, J. Yan, Y.-W. Chow, and M. Fahmideh, "A Comprehensive Overview of IoT-Based Federated Learning: Focusing on Client Selection Methods," *Sensors (Basel, Switzerland),* vol. 23, 2023-08-01 2023, doi: 10.3390/s23167235.

[47] J. Li, T. Chen, and S. Teng, "A comprehensive survey on client selection strategies in federated learning," *Comput. Networks,* vol. 251, p. 110663, 2024-07-01 2024, doi: 10.1016/j.comnet.2024.110663.

[48] M. Panigrahi, S. Bharti, and A. Sharma, "A review on client selection models in federated learning," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* vol. 13, 2023-09-04 2023, doi: 10.1002/widm.1514.

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 86-97
DOI: https://doi.org/10.25195/ijci.v51i2.628
Print ISSN: 2313-190X, Online ISSN: 2520-4912

[49] X. Wang, L. Ge, and G. Zhang, "A Review of Client Selection Mechanisms in Heterogeneous Federated Learning," pp. 761-772, 2023-01-01 2023, doi: 10.1007/978-981-99-4742-3_63.

[50] J. Kim, C. Song, J. Paek, J.-H. Kwon, and S. Cho, "A Review on Research Trends of Optimization for Client Selection in Federated Learning," *2024 International Conference on Information Networking (ICOIN),* pp. 287-289, 2024-01-17 2024, doi: 10.1109/ICOIN59985.2024.10572056.

[51] C. Smestad and J. Li, "A Systematic Literature Review on Client Selection in Federated Learning," *Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering,* 2023-06-08 2023, doi: 10.1145/3593434.3593438.

[52] D. B. Ami, K. Cohen, and Q. Zhao, "Client Selection for Generalization in Accelerated Federated Learning: A Multi-Armed Bandit Approach," *IEEE Access,* vol. 13, pp. 33697-33713, 2025, doi: 10.1109/ACCESS.2025.3543441.

[53] S. Moon and Y. Lim, "Client Selection for Federated Learning in Vehicular Edge Computing: A Deep Reinforcement Learning Approach," *IEEE Access,* vol. 12, pp. 131337-131348, 2024, doi: 10.1109/ACCESS.2024.3458991.

[54] S. Mayhoub and T. M. Shami, "A Review of Client Selection Methods in Federated Learning," *Archives of Computational Methods in Engineering,* vol. 31, no. 2, pp. 1129-1152, 2024/03/01 2024, doi: 10.1007/s11831-023-10011-4.