Review

Addressing Bias and Fairness using Fair Federated Learning: A Systematic Literature Review

Dohyoung Kim [*] , Hyekyung Woo [*] , Youngho Lee [*]

Posted Date: 15 October 2024

doi: 10.20944/preprints202410.1060.v1

Keywords: Distributed computing methodologies; Fair federated learning; Bias; Fairness; Privacy preservation

Review

Addressing Bias and Fairness Using Fair Federated Learning: A Systematic Literature Review

Dohyoung Kim [1], Hyekyung Woo [2,*] and Youngho Lee [3,*]

[1] Department of IT Convergence Engineering, Gachon University, Sungnam, Korea; rlaehgud9604@gachon.ac.kr
[2] Department of Health Administration, Kongju National University, Gongju, Korea
[3] Department of Computer Engineering, Gachon University, Sungnam, Korea
* Correspondence: hkwoo@kongju.ac.kr (H.W.); lyh@gachon.ac.kr (Y.L.)

Abstract: In the field of machine learning, the rapid development of data volume and variety requires ethical data utilization and strict privacy protection standards. Fair Federated Learning (FFL) has emerged as a key solution that aims to ensure fairness and privacy protection in a distributed learning environment. FFL enhances privacy protection and solves the inherent limitations of existing federated learning (FL) by promoting fair model training in diverse participant groups, preventing the exclusion of individual users or minorities, and improving overall model fairness. In this study, FFL discusses the causes of bias and fairness of existing FL, and separates solutions based on data partitioning strategies, privacy mechanisms, applicable machine learning models, communication architectures, and technologies to overcome heterogeneity. In order to improve the causes of bias, fairness, and privacy protection of FL, fairness evaluation indicators and applications and challenges of FFL are discussed. Since it addresses bias, fairness, and privacy issues in FL of all mechanisms, it can be an important resource for practitioners who want to implement efficient FL solutions.

Keywords: distributed computing methodologies; fair federated learning; bias; fairness; privacy preservation

1. Introduction

Due to the recent development of big data and LLM, issues of data, privacy protection, and security are being discussed more importantly than issues of data quantity. Representative examples include China's Cyber Security Law, the Civil Code of the People's Republic of China, the European Union's General Data Protection Regulation (GDPR), the U.S. Health Insurance Portability and Accountability Act (HIPAA), and Korea's Personal Information Protection Act, which focus on privacy protection and data security.

Centralized data processing, which is mainly used for artificial intelligence, has limitations such as privacy protection, data security, data silos and accessibility, bandwidth and efficiency, scalability, and bias. Therefore, Federated Learning (FL), which is a distributed learning method, allows individual users in different spatial locations to collaborate with other users to learn machine learning models compared to centralized learning methods, and all personal data that may contain sensitive personal information can be stored on the device. By solving these core problems, FL not only provides a practical solution to the limitations of centralized learning, but also opens up possibilities for AI applications in sensitive or restricted environments such as healthcare, finance, and mobile services, and can enjoy the advantage of obtaining well-trained machine learning models without having to send privacy-sensitive personal data to a central server.

However, despite the growing interest in FL as a privacy-preserving machine learning framework, it still has significant limitations in addressing inherent bias and fairness issues. Currently, FL has problems with bias and fairness due to the distributed nature of data sources, which must be solved to effectively protect personal information, such as personal information, insufficient data volume, and statistical heterogeneity. Since FL models may perform well in the majority group but may not accurately predict the results of the minority group due to differences in data distribution between clients, it becomes complicated to detect and mitigate bias, which leads to fairness issues. Bias in FL can arise from uneven data distribution across devices, and models may not perform equally well across all users or scenarios due to skewed local data representations. Similarly, fairness in FL has not been sufficiently explored, particularly in terms of how to ensure fair model performance across diverse participant groups with varying data quality and quantity. Addressing these issues requires sophisticated fairness-aware strategies that not only improve model accuracy across diverse datasets, but also ensure that no group or demographic is systematically disadvantaged. Thus, pursuing fairness in FL requires a delicate balance between protecting user privacy, ensuring robust model performance, and promoting fair outcomes across all participating entities.

The need for Fair Federated Learning (FFL) provides a mechanism to reduce the various problems of FL, especially imbalances such as data heterogeneity, unfair resource distribution among clients, and the problem of discriminatory model training. FFL proposes a methodology that not only respects data privacy but also actively mitigates bias and promotes equity in model results. This includes designing novel algorithms that can handle heterogeneous data distributions and promote fairness, and potentially tailoring the learning process to give more weight to disadvantaged groups or develop models that can generalize across diverse data sources. FFL incorporates specific strategies such as reweighting training data across multiple nodes, modifying learning objectives to emphasize fairness, and using regularization techniques to penalize unfair outcomes. In addition, FFL's privacy-enhancing strategies such as data distribution, application of encryption techniques, differential privacy, access control and authentication mechanisms, and transparency of the FL protocol help enterprises and institutions develop fair and unbiased AI models while protecting their customers' data. By incorporating these strategies, FFL maintains data in local nodes, thereby enhancing privacy, and significantly improving the fairness of the resulting model. This dual focus on privacy and fairness makes FFL a powerful framework, especially for applications such as healthcare, finance, and public services, where fairness and non-discrimination are as important as data security. Therefore, leveraging FFL focuses on improving the fairness of FL and ensuring that all participants benefit fairly. FFL’s data privacy, model fairness, and non-discrimination effectively address the ethical issues inherent in machine learning systems by addressing privacy and fairness issues that may arise in centralized learning systems. This FFL approach can contribute to making FL a more practical and sustainable technology by addressing both technical and ethical issues. The evolution to FFL is essential for deploying FL in sectors such as healthcare and banking, where decision-making impacts are profound and have a direct impact on human lives.

This study focuses on systematically exploring how existing research addresses (or fails to address) the nuances of bias and fairness in the FL environment. We emphasize the need to integrate fairness into the FL paradigm and describe potential research directions for improving the reliability and applicability of FL models to ensure fairness and privacy. It is meaningful to study how technologies that address these bias and fairness issues can be further developed in the future. In addition, we systematically review recent research on privacy and data security and review, and examine the challenges and directions that should be derived in the future. The main contributions of this study are as follows: (1) We reviewed the development of FFL. (2) We explained the paper preprocessing process. (3) We reviewed the sources of bias and fairness in FL. (4) We reviewed the work of FFL in three aspects: data segmentation, privacy protection mechanisms, and methods to address heterogeneity. (5) We discussed the applications, challenges, and future research directions of FFL.

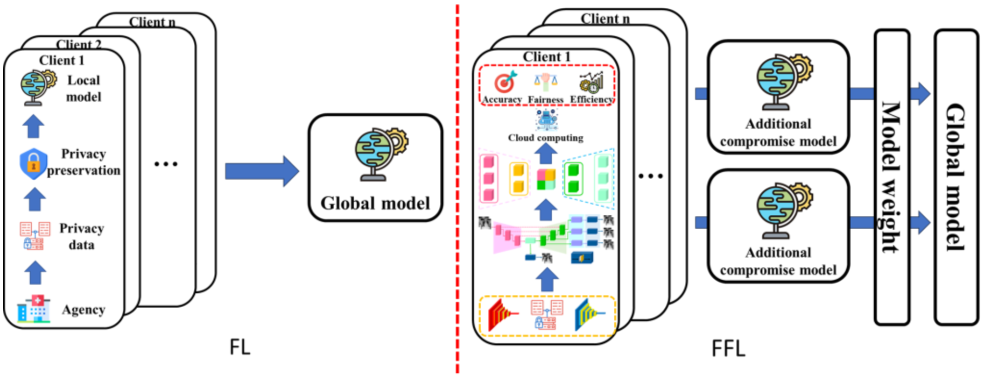

Figure 1. Overview of FL, FFL.

2. Methodology

This study followed the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)" guidelines[1].

2.1. Literature Review and Search Strategy

We performed an extensive search of articles published up to July 2024 in databases such as PubMed, Embase, Web of Science, IEEE Xplore Digital Library, ACM Digital Library, ELSEVIER, SpringerLink, Google Scholar, Semantic Scholar, Cornell University, Computing Research Repository, Database Systems and Logic Programming, and identified relevant English-language articles using the following search strings, slightly modified to fit the database: "Fair Federated Learning", "Fairness-aware Federated Learning", "Incentive Federated Learning", "Efficient Federated Learning", "Privacy-preservation Federated Learning", "Recent Advanced Federated Learning", "Fair Distributed computing methodology". Similarly, we performed searches in Chinese databases CNKI and Wanfang. In addition, we manually searched various commentaries, editorials, and conference proceedings to ensure completeness.

2.2. Study Selection

Studies were included if they met the following criteria: 1) Bias of Federated Learning; 2) Problem of Federated Learning; 3) Future work of Federated Learning. Case reports, post hoc analyses, reviews, meta-analyses, and studies published in languages other than English and Chinese were excluded.

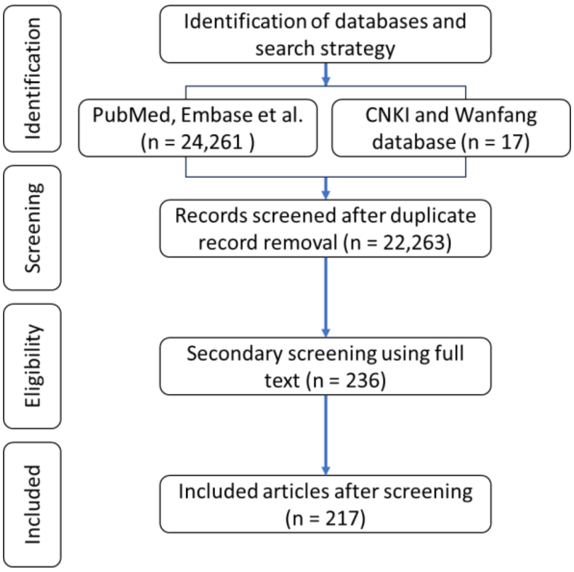
Identification
Screening
Eligibility
Included
Identification of databases and search strategy
PubMed, Embase et al. (n = 24,261)
CNKI and Wanfang database (n = 17)
Records screened after duplicate record removal (n = 22,263)
Secondary screening using full text (n = 236)
Included articles after screening (n = 217)

Figure 2. Preferred Reporting Items for Systematic reviews and Meta-Analysis diagram.

2.3. *Selection of Eligible Articles*

The initial electronic search yielded 24,278 integrated studies from the selected databases, and after removing duplicates, 22,263 articles remained. Articles were further reviewed for eligibility based on titles and abstracts. After full-text review of these articles, 217 articles were deemed suitable for the final analysis. Figure 2 shows the study flow diagram according to the PRISMA guidelines. One Chinese article and 216 English articles were included in the final analysis. All studies were published between 2015 and 2024. Figure 3 shows the trends in FFL research over the past five years as investigated in this study. Figure 4 shows the default settings for each FFL item.

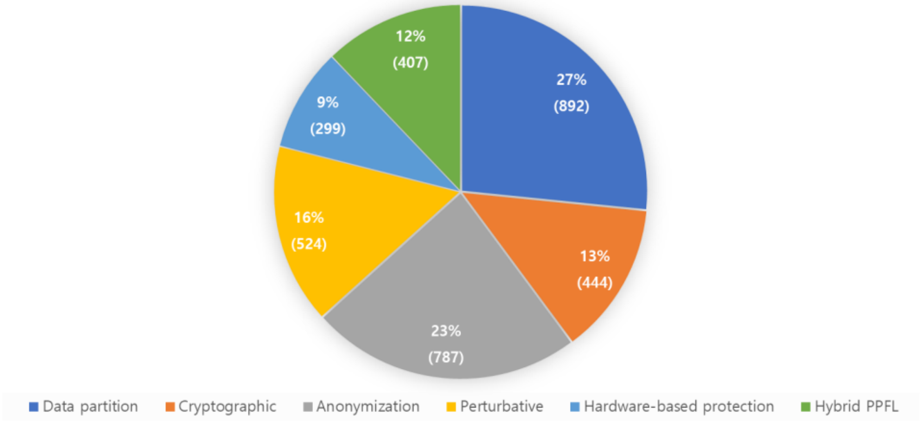

Figure 3. Trends in FFL research over the latest 5 years.

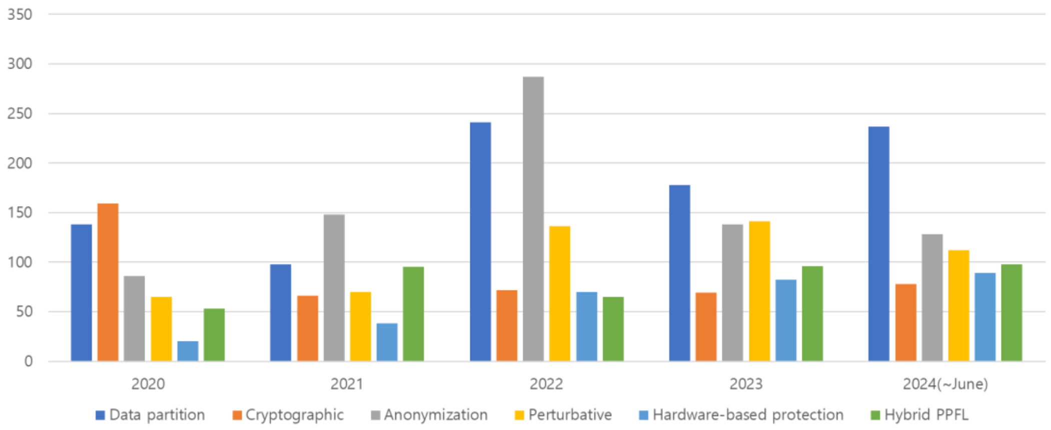

Figure 4. Preference status by FFL category.

3. Causes of bias in FL and fairness in FFL

The systematic review covered a wide range of studies focusing on the causes of FL bias by category, as detailed in Table 1. The main findings of the study highlight bias and fairness in improving privacy in FL. The analysis was subdivided into several predefined criteria: conventional bias sources, sub-sampling, party selection, dropouts, heterogeneity, and method. Figure 1 shows the framework of conventional FL and FFL.

Table 1. Systematic review of bias in FL.

Categorizations	Ref.	Methods	Main conclusions
Conventional bias sources	[2]	Bias propagation	Analyzing the bias propagation of FL on real-world datasets, we show that biased parties unintentionally covertly encode bias in a small number of model parameters, steadily increasing the global model's reliance on sensitive attributes throughout training.
	[3]	DCFair	Focusing on the impact of bias, we explore how factors such as device interruptions, biased device data, and biased participation affect the FL process from a system perspective. We present a characterization study that empirically demonstrates how these challenges affect important performance metrics such as model error, fairness, cost, and training time, and why it is important to consider them together rather than individually, and describe a method called DCFair, a framework that comprehensively considers several important challenges of real-world FL systems.
	[4]	FFL	We discussed how to eradicate the source of bias and create a more fair federated learning environment.
Sub-sampling, party selection and dropouts	[5]	FL-FDMS	Based on convergence analysis, we develop a novel algorithm, FL-FDMS, which discovers a client's friends (i.e., clients with similar data distributions) on the fly and uses the friends' local updates as replacements for dropout clients to reduce errors and improve convergence performance.
	[6]	MimiC	The MimiC algorithm server modifies each received model update based on the previous update. The proposed modification to the received model update mimics a virtual central update regardless of the interrupted client. Theoretical analysis of MimiC shows that the difference between the aggregated update and the central update is reduced with an appropriate learning rate, leading to convergence.
	[7]	FedDebias	To address the unexplored phenomenon of biased local learning that can explain the problems caused by local updates in supervised FL, we propose FedDebias, a novel integrated algorithm that reduces the local learning bias of features and classifiers.
Data heterogeneity	[8]	FL solutions	We consider the problem of fault identification and simulate various data heterogeneity scenarios, demonstrating that these settings remain challenging for popular FL algorithms and should be taken into consideration when designing a federated predictive maintenance solution.
	[9]	FedNH	We initially showed that uniformly distributing class prototypes across the latent space and smoothly injecting class semantics into class prototypes and enforcing uniformity helped prevent prototype collapse, and that injecting class semantics improved the local model.

	[10]	non-IID settings	By analyzing the principal angles between subspaces in different classes of each dataset, we propose a new concept and framework for non-IID segmentation in FL settings.
Fusion methodologies	[11]	FedDRL	Including malicious models in the fusion process that uses weighted average techniques for model fusion can significantly reduce the accuracy of the aggregated global model, and to address the problem of FL where the number of client samples does not determine the weight values of the model due to the heterogeneity of devices and data, we propose a reliable model fusion method based on reinforcement learning (FedDRL).
	[12]	Aggregation method	We investigated various aggregation methods that could affect the fairness and bias of the resulting model.
Systems heterogeneity	[13]	Challenges of FL	We discuss the unique characteristics and challenges of FL, provide a broad overview of current approaches, and suggest several directions for future work relevant to the broader research community.
	[14]	HFL	We summarize the various research challenges of HFL from five aspects: statistical heterogeneity, model heterogeneity, communication heterogeneity, device heterogeneity, and additional challenges. We also review recent progress in HFL, propose a new taxonomy of existing HFL methods, and deeply analyze their advantages and disadvantages.
	[15]	Client selection mechanism	We propose a client selection mechanism that considers both system and statistical heterogeneity, aiming to improve the time-to-accuracy performance by offsetting the impact of system performance differences and inter-client data distribution differences on training efficiency.

3.1. Causes of Bias in FL

The causes of bias in FL are generally as follows:

3.1.1. Conventional Bias Sources

Conventional Bias Sources in FL come from various causes that can occur just like in traditional machine learning environments. These traditional sources of bias are not unique to FL, but are connected to problems found in general data science and AI systems. Historical Bias, which may reflect the way data was collected or the real-world inequalities or stereotypes represented by the data, is a bias inherent in the data itself[16].

Measurement Bias occurs due to errors in the data collection process[12]. Sampling Bias occurs when the dataset does not adequately represent the entire population[17]. If data from a specific group is oversampled or omitted, this can bias the model's ability to generalize. Aggregation Bias can occur during the process of aggregating models learned from multiple clients[18]. If the data distribution of each client is different, data from a specific client may have a greater impact than data from other clients during the central aggregation process. These traditional sources of bias should be considered in the design of FL systems, and if not properly managed through data collection, processing, and model learning methods, they can result in unfair predictions or decisions.

3.1.2. Sub-Sampling, Party Selection, and Dropouts

Among biases, subsampling, party selection, and dropout are particularly noteworthy due to their impact on the diversity and representativeness of data used to train models. The training purpose of subsampling refers to the practice of selecting a subset of data from a larger dataset. If the subset is not representative of the entire dataset, it can introduce bias and unintentionally emphasize or underestimate certain patterns or features in the data, which can distort the learning results.

Party selection in FL involves selecting clients (or parties) to contribute to the training of a shared model. The selection process can be a source of bias if certain types of clients are consistently selected over others. This can result in the model not performing well for all potential users, especially those whose data profiles are not well represented during the training process. Dropout refers to a situation where a client unexpectedly fails to complete a training task or fails to send updates back to the central server or aggregator of FL. This can lead to incomplete learning cycles, and the aggregated model updates can be distorted due to missing model updates from some clients. Frequent dropouts can lead to a type of selection bias, where the model overfits to the characteristics of clients with more stable or reliable engagement.

3.1.3. Data Heterogeneity

In FL, data heterogeneity refers to the variation in data distribution among different clients participating in the learning process. Heterogeneity is one of the major challenges of FL because it can significantly affect the performance and fairness of the resulting model. In FL, data heterogeneity occurs when data from different clients or devices differ in terms of distribution, volume, and features. This can occur due to different environments, user behaviors, and data collection protocols of different devices.

The impact of data heterogeneity on FL includes (1) Model performance: A model trained in a heterogeneous data environment may have difficulty generalizing well across all clients. A model that performs well on one client may not perform well on other clients due to differences in data characteristics. (2) Convergence speed: Data heterogeneity can slow down the convergence of the FL process. (3) Fairness issues: When data is heterogeneous, the federated model may be unintentionally biased toward data patterns of more dominant or more frequently participating clients.

3.1.4. Conventional Bias Sources

The impact of Fusion Methodologies on bias and fairness includes: (1) Variance reduction: A properly designed fusion methodology can reduce the variance of model performance across clients, enabling more stable and reliable predictions. (2) Fairness improvement: By adjusting the way local models are aggregated, biases that may arise from imbalanced or unrepresentative local datasets can be mitigated. (3) Generalization: Effective fusion can leverage the diversity of local datasets to improve the generalization of global models to unseen data and clients. Therefore, future directions for Fusion Methodologies focus on developing more adaptive, resilient, and fair methods for aggregating local models. This includes hybrid approaches that combine multiple methods and apply advanced machine learning techniques such as meta-learning and reinforcement learning to optimize the fusion process.

3.1.5. Systems Heterogeneity

System heterogeneity refers to the variability in hardware, network connectivity, computational power, and storage capacity of devices participating in the FL network. This type of heterogeneity is important because it directly affects the efficiency and effectiveness of the learning process across the distributed network. The effects of system heterogeneity on FL include: (1) Model learning efficiency: System heterogeneity can lead to significant differences in the time it takes for different devices to complete their respective share of the learning task. (2) Bias and fairness: If not properly managed, system heterogeneity can lead to bias in the training process.

3.2. Fairness and Recent Trends in FFL Research

FFL is a technique designed to provide a fair learning environment among various clients. This fairness is mainly classified into four types: Performance Fairness, Collaboration Fairness, Model Fairness, and Different client types Fairness. Each type plays an important role in the FFL environment and aims for balanced model development among clients.

Performance Fairness is fairness that ensures that all participating clients experience similar levels of model performance and is based on the principle that all participants should receive equal benefits. Since learning outcomes may vary depending on the amount or quality of data held by individual clients, performance fairness ensures that all participants contribute equally and receive equal value. Performance fairness increases motivation among participating clients and improves the reliability and satisfaction of the entire system.

Collaboration Fairness is fairness that ensures that all clients participate fairly in the learning process and that the costs and benefits incurred in the process are appropriately distributed. Collaborative fairness is important when multiple clients have different resources and helps all participants to be treated fairly and maintain participation. This promotes cooperation among participants and encourages long-term project participation, ensuring the performance and stability of the entire network.

Model Fairness refers to the fairness with which the trained model shows equal performance for various demographic groups. Model fairness is necessary to train a model without bias toward a specific group. A fairly designed model respects the diversity of users and becomes competitive in the market. Each of these types of fairness is essential to ensure that all participants benefit equally and contribute fairly in the complex environment of FFL. This enables the construction of a more sustainable and inclusive AI system.

Different client types Fairness considers the unique differences between clients, such as different computational capacities, dataset sizes, or data distributions. To fairly accommodate these differences, the FL process is customized to ensure that no client is disadvantaged due to its unique characteristics. The various fairness methods possessed by FFL are shown in Figure 5.

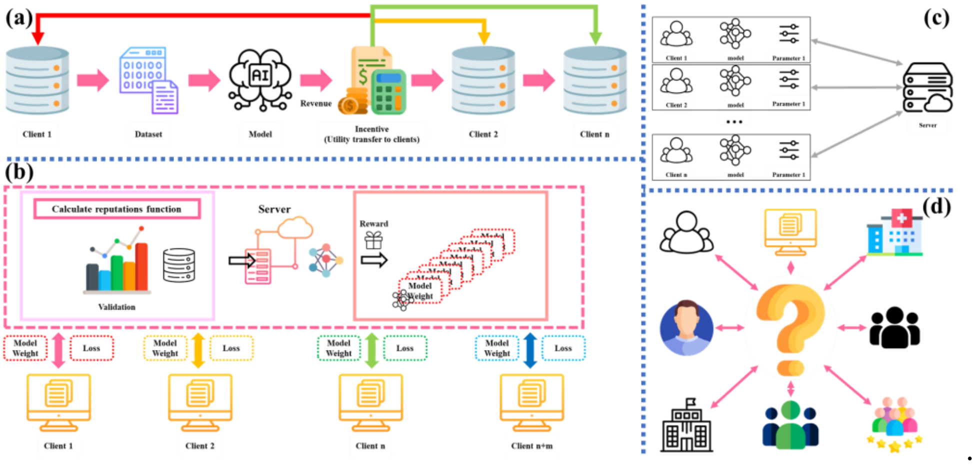

Figure 5. Various types of fairness methods in FFL. (a) Performance fairness such as accuracy parity, (b) Collaborative fairness, (c) Model fairness, (d) Different client type fairness.

The purple mark on the map in Figure 6 is the distribution of FFL authors. Countries that received scores higher than 5 were the United States (80), China (78), Australia (11), Singapore (10), Republic of Korea (9), Canada (8), Great Britain (8), France (8), Israel (1), Sweden (1), Türkiye (1) Italy (1), and Denmark (1).

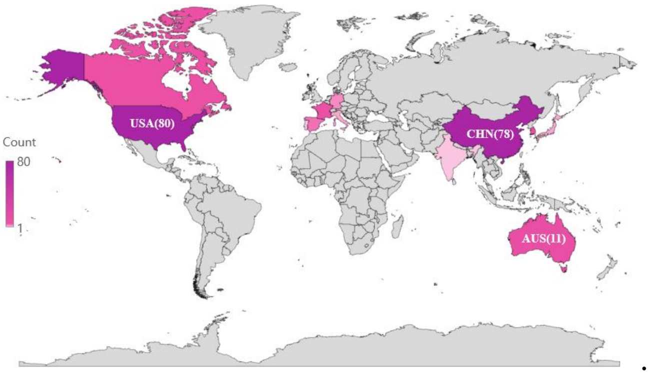

Figure 6. Geographical distribution of FFL authors.

4. Bias for categorizations

In this section, we summarize the classification of FFLs in terms of five aspects: data partitioning, privacy mechanisms, applicable machine learning models, communication architectures, and heterogeneity resolution methods. The FFL metrics are listed in Table 2, and the fairness concept is illustrated in Figure 8.

Table 2. Categorizations of FFL.

Categorization	Methods	Algorithm	Advantage	Applications
Data partitioning	Fair horizontal federated learning (FHFL)	FFL-OppoGAN[19], FedUB[20]	Increases user sample size	Android phone model updates; logistic regression
	Fair vertical federated learning (FVFL)	FairVFL[21], vflow[22], MOSP[23], FedOnce[24], FedUFO[25], UAB[26], BadVFL[27]	Increases feature dimension	Decision trees; neural networks
	Fair federated transfer learning (FFTL)	FAIR-FATE[28], FCFL[29], DRFL[30], TFCS[31]	Increases user sample size and feature dimension	Transfer learning
	Multimodal fair federated learning(MMFFL)	FedUFO[25], FedCMI[32], mFairFL[33], MAFL[29], PFL-MCL[34]	Enhances data security, powerful performance, accessibility, scalability, and data use	Securing multi-dimensional fairness, optimizing integrated model learning processes,

				metaverse applications
Privacy mechanism	Homomorphic encryption (HE)	CKKS[35], FV[36], FLASHE[37], FedML-HE[38]	Users can calculate and process encrypted data	Ridge regression; federated learning
	Secure multi-party computation (SMPC)	Obliv-C[39], chain-PPFL[40], FL rack[41], VFL-R[42], SecureML[43], PrivFairFL[44], SMPAI[45], SMPC[46]	Minimizes risk of information leakage, optimizes computation across multiple participants	Effective scaling to large networks
	Secret sharing (SS)	HFTL[47], PFK-means[48], VerifyNet[49], FairFed[50]	Strengthens security, elasticity, and flexibility	Distributes important information among managers or systems
	Data masking (DM)	FedMask[51], PerFedMask[52], FedLMD[53], GlueFL[54], PrivMaskFL[55]	Trains models without exposing local data	Develops a recommendation system based on user behavior while protecting user identity
	Data shuffling (DS)	CLDP-SGD[56], SS-Double[57], MSFL[58]	Resolves data imbalance and prevents overfitting	Data characteristics are evenly reflected in the model learning process
	Differential privacy (DP)	DPBalance[59], mFairFL[33], FFL-OppoGAN[19], FairFed[60], Local DP FL[61], PEFL[62], OARF[63], FL-LSTM[64], Client-Level DP FL[65], Hybrid	Protects user privacy by adding noise	Conventional machine learning; deep learning

	FL[66], FedFDP[67], DP-DLP[68]		
Additive perturbation methods (APM), Multiplicative perturbation methods (MPM)	FedISM[69], PILE[70], Fed-SMP[71], ANP[72], DISTPAB[73]	Maintains data usefulness while enhancing data privacy, Protects data privacy while preserving original data structure and relationships	Incorporates random noise into data with a specific distribution, Prevents exposure of personal identity or sensitive information; ensures compliance with data protection regulations (e.g., GDPR)
Hardware-based protection	FFL-OppoGAN[19], FairFed[60], ShuffleFL[74], FLASH[75], FLATEE[76], EdgeFed[77]	HSM and TPM protect against hardware physical attacks	Hardware solutions often provide faster throughput than software alone
Hybrid privacy-preserving federated learning (HPPFL)	RVE-PFL[78], HybridAlpha[79], xMK-CKKS[80], LEGATO[81], APPFed[82], FederatedTrust[83], PrivFairFL[44], FPFL[84], HFAD[85]	Simultaneously provides improved data privacy and model efficiency	Adapts protection techniques to different environments and requirements; improves efficiency of the entire system
Model aggregation (MA)	PATE[86]	Avoids transmitting original data	Deep network federated learning; PATE method

	Sensitivity-based weight distribution	FFLFCN[87], FedSW[88]	Improves accuracy, efficiency, and fairness	Healthcare, finance, smart cities
Applicable machine learning model	Linear models	GLM[89], FedUFO[25], HDP-FL[90], FairFed[50], FFM[91], mFairFL[33]	Concise format, easy to model	Linear regression; ridge regression
	Tree models	q-FFL[92], FedStaleWeight[93], FAIR-FATE[28], FedFaiREE[94]	Accurate, stable, and able to map non-linear relationships	Classification trees; regression trees
	Neural network models	RSRA[95], SpreadGNN[96]	Exhibits learning capabilities, high robustness, and fault tolerance	Pattern recognition, intelligent control
Heterogeneity resolution methods	Asynchronous communication(AC)	Asynchronous FL[97], FedBuff[98], ASO-Fed[99], AFAFed[100]	Prevents communication delay	Device heterogeneity
	Sampling	CFFL[101], FLI[102], Fold-stratified cross-validation[103], FedSSAR[104], ISFL[105], DivFL[106], Delta[107]	Avoids simultaneous training with heterogeneous equipment	Pulling reduction with local compensation (PRLC)
	Fault-tolerant mechanism	FEEL[108], BDFL[109]	Prevents entire system from collapsing	Redundancy algorithm
	Heterogeneous model	FCCL[110], FedAlign[111], FedRolex[112]	Solves corresponding heterogeneous device	LG-FEDAVG algorithm

4.1. Data Partition

FFL's data partitioning method is a differentiated approach that takes the fairness of FL as an important goal. As shown in Figure 7, according to the different distribution patterns of the data sample space and the feature space, FFL is divided into four categories: fair horizontal federated learning, fair vertical federated learning, fair federated transfer learning, and multimodal fair federated learning.

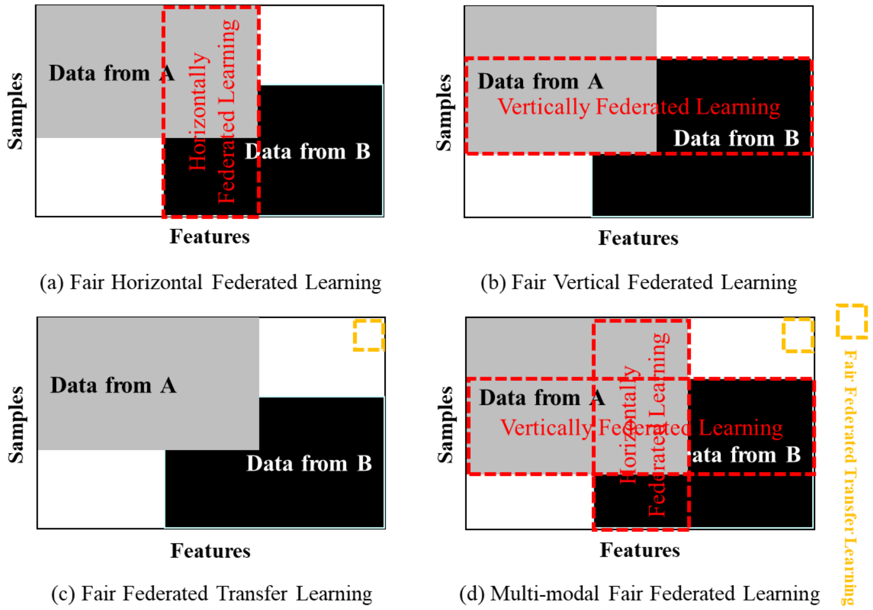
Samples
Data from A
Horizontally Federated Learning
Data from B
Features
(a) Fair Horizontal Federated Learning
Samples
Data from A
Vertically Federated Learning
Data from B
Features
(b) Fair Vertical Federated Learning
Samples
Data from A
Data from B
Features
(c) Fair Federated Transfer Learning
Samples
Data from A
Vertically Federated Learning
Horizontally Federated Learning
ata from B
Features
(d) Multi-modal Fair Federated Learning
Fair Federated Transfer Learning

Figure 7. The different data partition of fair horizontal federated learning, fair vertical federated learning, fair federated transfer learning, and multi-modal fair federated learning.

4.1.1. Fair Horizontal Federated Learning (FHFL)

Horizontal Federated Learning (HFL) is designed for scenarios where multiple participants (clients) share the same feature space but have different datasets with different samples. However, this form of partitioning inherently carries the risk of bias and unfairness. Different data such as data heterogeneity, imbalanced data, and non-IID data can lead to poor performance or bias due to distorted model training. FHFL is improved through various approaches that aim to mitigate bias and ensure more fair results for all participants, and ensure that the FL process is sensitive to data distribution and demographic differences among participants.

In FL, the bias is higher than that of centralized learning using models trained on the union of all data (1). For a wide range of fairness metrics, a global fairness training algorithm that can obtain the fairness of the global model using only summary statistics of local clients and directly minimizes the empirical loss was proposed[113]. An adversarial training method was proposed to train a FL model that satisfies individual fairness without accessing the underlying data distribution, thereby improving individual fairness and group fairness and protecting privacy[114].

4.1.2. Fair Vertical Federated Learning (FVFL)

Vertical Federated Learning (VFL) can cause bias in the learning process due to data imbalance, feature mismatch, privacy requirements, etc., which can reduce accuracy. Therefore, in order to solve the bias problem in FVFL, Secure Multi-Party Computation (SMPC), Fairness Algorithms, balanced data sampling, and weight adjustment can reduce bias while protecting privacy.

4.1.3. Conventional Bias Sources

Federated Transfer Learning (FTL) suffers from bias due to the lack of data homogeneity and limitations of pre-trained models when the model over-reflects the characteristics of the pre-trained data and fails to generalize to new data. Therefore, FFTL generalizes the model by using a multivariate approach, data compensation techniques, and adaptive learning processes to reduce bias and better adapt to new environments. Coston et al. proposed a new weighting method called Prevalence-Constrained Covariate Shifting (PCCS), which does not require protected attributes in the target domain, and Target-Processed Covariate Shifting (TFCS), which does not require protected attributes in the source domain[31].

4.1.4. Conventional Bias Sources

Multi-modal Federated Learning (MMFL) is a technology that learns by integrating various types of data modes (images, text, audio, etc.), but there is a risk of model bias due to imbalances between each data mode or overrepresentation of specific modes, which may reduce generalization ability. To address this bias, MMFFL adjusts weights through data balancing techniques, algorithm modifications, and diversification of model evaluations, and alleviates bias through independent evaluation metrics.

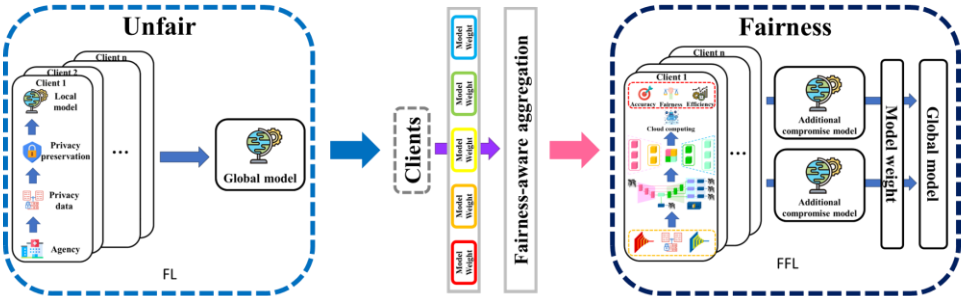

Figure 8. Overview of fairness in FFL.

4.2. Privacy Preservation Mechanism

The main privacy protection mechanisms used in FFL include encryption methods, anonymization methods, perturbation methods, hardware-based protection, and Hybrid PPFL methods. These are similar to the mechanisms of basic FL, but FFL places greater emphasis on fairness and privacy. Figure 9 shows various types of privacy protection methods in FFL. Personal information protection indicators include pseudonymization technology and personal information protection technology.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 October 2024 doi:10.20944/preprints202410.1060.v1

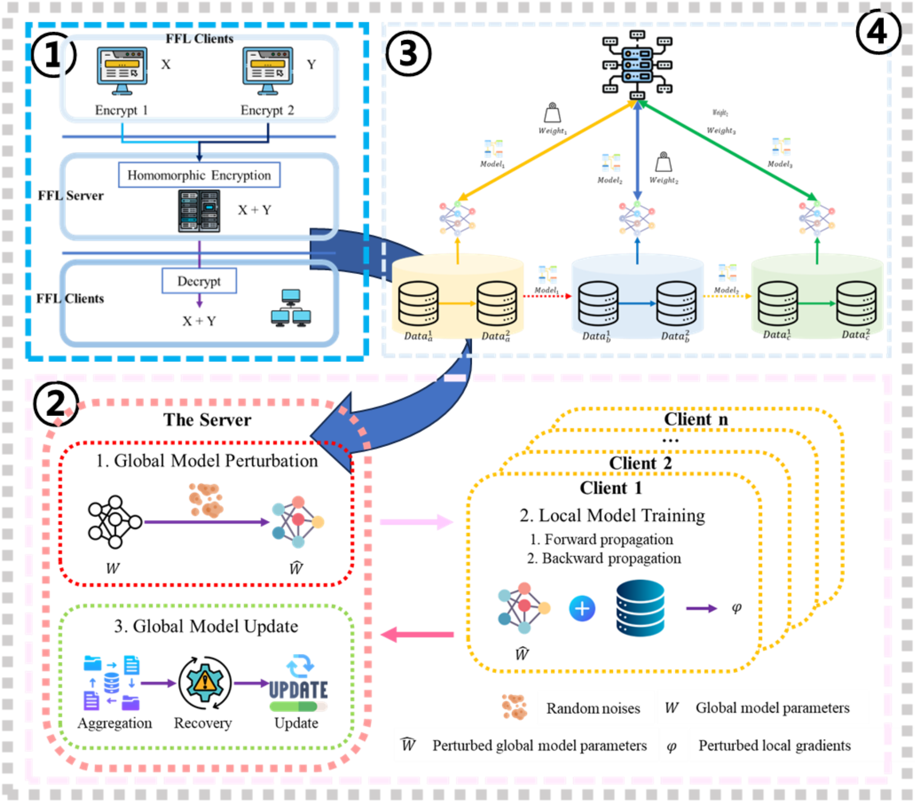

Figure 9. Various types of privacy preserving methods in FFL. ① Cryptographic method based on homomorphic encryption, ② Anonymization method, ③ Perturbation method, ④ Hybrid privacy-preserving method will be a collaboration of the three methodologies.

Pseudonymization technology is used to maintain the usability of data while minimizing the possibility of personal identification. It is a method of lowering the possibility of identifying original data by removing or replacing direct identifiers from personal information. Table 3 presents the current landscape of personal information pseudonymization technologies.

Table 3. Comprehensive landscape of pseudonym processing technologies.

Categorization **Privacy mechanism**	**Personal Information protection commissioner's pseudonymous information processing guidelines[115]**	**Guidelines for processing pseudonymized information in the education field[115]**	**Guide to pseudonyms and anonymization in the financial sector[115]**	**Guidelines for using healthcare data[116]**
Deletion technology	Deletion, partial deletion, line-item deletion, local deletion, masking	Record deletion, column deletion, partial deletion, deletion of all identifying elements	Masking, local deletion, record deletion	Identifier: deleted or replaced with serial number
Statistical tools	Total processing, partial processing	Sampling, total processing	Sampling, total processing	Key personal information: Reduc

Generalization skills	General rounding, random rounding, control rounding, top and bottom coding, local generalization, range method, problem data categorization	Rounding, general rounding, random rounding, control rounding, local generalization, top and bottom coding, attribute combination (categorization)	Rounding, top and bottom coding, combining attribute sets into a single attribute value (categorization), local generalization	e identification by deleting or extracting some meaningful information Attribute value: Apply pseudonymization techniques for each data type, such as deletion and masking
Encryption technique	Bidirectional encryption, one-way encryption, order-preserving encryption, form-preserving encryption, homomorphic encryption, polymorphic encryption	Deterministic encryption, order-preserving encryption, form-preserving encryption, homomorphic encryption, homomorphic secret distribution	Deterministic encryption, order-preserving encryption, form-preserving encryption, homomorphic encryption, homomorphic secret distribution	Some data types do not require pseudonymization (e.g., measurement value information) and pseudonymization is limited (e.g., voice data, biometric data, and genome data)
Randomization technique	Noise addition, permutation (permutation), tokenization, random number generation	Add noise, permutations, 1:1 swaps, partial sums	Permutations, noise addition, partial totals	
Other technologies	Sampling, dissection, data reproduction, homomorphic secret distribution, differential privacy	Anatomy, data reproduction	Anatomy, data reproduction	
Pseudonymization technique	No relevant technique or classification criteria available	Mapping table, counting, pseudorandom number generation, hash algorithm, bidirectional encryption, masking, scrambling, blurring, token system, polymorphic encryption	Mapping table, bidirectional encryption, one-way encryption, tokenization	

Privacy protection model	No relevant technique or classification criteria available	No relevant technique or classification criteria available	k-anonymity, l-diversity model, t-proximity model, differential privacy, protection model	

4.2.1. Fair Cryptographic methods

4.2.1.1. Homomorphic Encryption, HE

In FFL, HE is used to ensure that each participant's data is processed fairly, and analysis can be performed without the risk of sensitive information leakage. Among the current HE methods, Fully Homomorphic Encryption is relatively more efficient and supports larger ciphertext sizes, but it is a slow process due to the large amount of computation required, while Partially Homomorphic Encryption is less efficient and supports smaller ciphertext sizes, but it is a fast process and requires less computational power. Jin et al. proposed FedML-HE, the first practical FL system with efficient HE-based security model aggregation, which significantly reduces computation and communication overhead during learning by selectively encrypting sensitive parameters, while providing customizable privacy features[38].

4.2.1.2. Secure Multi-party Computation, SMPC

In FFL, SMPC guarantees the data privacy of participants by performing necessary computation without directly sharing data, and optimizes computation among multiple participants through multi-party computation, which can effectively scale even in large-scale networks. PrivFairFL by Pentyala et al. proposed a method to train group fair ML models in privacy-preserving and cross-device FL without requiring clients to disclose sensitive attribute values by combining FL with SMPC and DP[44]. SMPAI by Mugunthan et al. reduces privacy and data leakage risks based on SMPC and DP[45]. FL-IPFE by Chen et al. proposed a new SMPC algorithm to protect local gradients[46].

4.2.1.3. Secret Sharing, SS

As data security and privacy become increasingly important in FFL, secret sharing has become an essential tool for safely storing and managing access to critical data. This is especially necessary for distributing critical information among various managers or systems. FairFed by Ezzeldin et al. enabled group fairness using secret sharing techniques in a FL setting[50].

4.2.2. Anonymization methods

Data masking and shuffling are attempts to preserve data validity as much as possible while maintaining data confidentiality by transforming the actual value of the data or changing the order, respectively.

4.2.2.1. Data Masking

In FL, data masking is useful for enabling each client to collaborate with the central server to train a model without directly revealing its local data. Each client can participate in model training without exposing its personal information by masking its own data and sending it to the central server. The PrivMaskFL framework by Xiong et al. uses dynamic participant aggregate masking together with adaptive differential privacy to address bias and fairness issues and reduce communication overhead in FL environments[55].

doi:10.20944/preprints202410.1060.v1

4.2.2.2. Data Shuffling

In FL, data shuffling can resolve data imbalances between clients and prevent models from being overfitted to the data distribution of a specific client. When training a disease prediction model using medical data from various regions, the characteristics of the data collected from each region are different, so shuffling can ensure that the data characteristics of each region are evenly reflected in the model training process.

4.2.3. Perturbative methods

4.2.3.1. Differential Privacy, DP

DP in the existing FL is a privacy-preserving technology that adds small noise to the query response to prevent personal information from being inferred. In FFL, DP focuses on maintaining the accuracy of the overall model while preventing the exposure of each participant's data, and helps reduce discrimination that may arise from data-based decisions of individual participants, so that all participants are treated fairly. Ling et al.'s FedFDP used DP to adjust fairness while maintaining privacy and incorporated clipping applied to the loss value to maximize utility[67]. Salim et al. developed a DP-DLP mechanism to properly adjust the extended variance of Gaussian noise to meet the requirements of DP at a unique privacy level and hide the influence of a single data point on the learning stage[68].

4.2.3.2. Additive/Multiplicative Perturbation methods

The greatest advantage of additive perturbation is its ability to preserve the usability of data while enhancing privacy[117]. It protects privacy by adding controlled noise to data, while at the same time preserving the core statistical properties of data required for data analysis or machine learning algorithms. The greatest advantage of multiplicative perturbation is that it can preserve the privacy of data while preserving the original structure and relationships of data[118]. It effectively reduces the sensitivity of data while preserving the usability of data as much as possible. As a result, it can maintain high accuracy in data analysis and machine learning tasks.

PILE by Tang et al. developed a verifiable perturbation scheme that makes confidential local gradients verifiable for gradient verification[70]. Hu et al. proposed Fed-SMP, which mitigates the impact of privacy on model accuracy by using SMP where the local model is first sparse before being perturbed by Gaussian noise[71]. Li et al. proposed ANP, a stepwise noise perturbation operation, to better balance privacy and accuracy[72].

4.2.4. Hardware-based protection

Hardware-based protection methods in FL are approaches that use specially designed physical hardware devices to enhance the security and privacy of data. Hardware-based methods often provide stronger security than software-based methods, making them more resistant to a variety of cyber threats. Researchers at Imperial College London are exploring the intersection of systems, security, and machine learning, focusing on using hardware hardening mechanisms to ensure the confidentiality and integrity of FL models, helping to mitigate risks associated with bias and ensure fairness in the training and deployment of models[119].

4.2.5. Hybrid Privacy Preserving Federated Learning, Hybrid PPFL

Hybrid PPFL method in FL aims to protect data more effectively by utilizing the strengths of each privacy technology and complementing its weaknesses. For example, combining encryption and DP can improve data security and usability. It can complement vulnerabilities that may occur when using only one privacy protection mechanism and performance degradation issues by combining them in various ways, and can adjust protection techniques to suit various environments and requirements.

4.2.5.1. Model Aggregation

In the area of FFL, model aggregation has been studied to mitigate bias and fairness issues, especially focusing on fair and privacy-preserving methods. In model aggregation, privacy issues are addressed while utilizing techniques such as SMPC for effective model aggregation. In FFL, additional weights or adjustments can be made to ensure fairness of each participant during the model aggregation process.

Guo et al. used encryption and hybrid models to enhance both privacy and fairness in FL systems[119]. They utilize confidential computing mode and secure encryption to isolate data and model learning processes from potential interference, thereby maintaining the fairness of model results. Truex et al. presented a hybrid model that combines various privacy mechanisms to ensure both data protection and fairness in model aggregation[120].

4.2.5.2. Sensitivity-based Weight Distribution

Sensitivity-based weight distribution is a hybrid approach that can help optimize model learning efficiency while maintaining data privacy. It ensures that all data sources contribute fairly to model learning, controls sensitive data from excessively influencing the model, and distributes weights according to importance to optimize learning efficiency.

Chatzikonstantinou et al. proposed a balanced and fair model by implementing a weight distribution technique that can effectively manage the difference in data contributions from various clients[121]. Dilley et al. modified the FedAvg method to introduce fairness constraints to ensure fair model training on diverse client datasets[122]. Chen et al. maintained local consistency while ensuring global convergence, reduced bias, and improved fairness in diverse data domains[123]. Zhao et al. effectively addressed bias by compensating for uneven data quality and distribution of clients[88].

4.3. Methods for solving heterogeneity

Methods for resolving heterogeneity in FFL are similar in basic structure to basic FL methods, but provide a differentiated approach, especially in terms of enhancing fairness and efficiency. These methods focus on asynchronous communication, sampling, fault tolerance mechanisms, and model heterogeneity. These methods effectively manage the diversity and heterogeneity of data through FFL and contribute to creating an environment in which all participants are treated fairly.

4.3.1. Asynchronous Communication

In the existing FL environment, not all clients may have the same network status or computational power. In FFL, asynchronous communication is further advanced to determine the priority of each client, ensuring more fair resource distribution and participation opportunities, and especially improving efficiency in environments where communication costs are high.

Chen et al. proposed an asynchronous online federated learning ASO-Fed framework in which edge devices perform online learning with continuous streaming local data and a central server aggregates model parameters of clients[124]. Baccarelli et al. proposed a novel asynchronous fair adaptation FL framework for stream-oriented IoT application environments, featuring time-varying operating conditions, heterogeneous resource constraints, non-IID local training data, and unreliable communication links[100]. Chen et al. proposed a novel easy-to-implement FL algorithm that controls the mismatch between the global model and the localized model using asynchronous settings and strategies, and adjusts the number of local epochs by estimating the aging, accelerating convergence, and preventing performance degradation due to stragglers[125]. Wang et al. proposed a novel asynchronous FL framework that adapts well to the heterogeneity of users, communication environments, and learning problems by considering both the potential delay in learning and uploading local models and the resulting staleness among receiving models that significantly impacts global model convergence[124].

4.3.2. Sampling

In FFL, sampling includes data imbalance resolution, stratified sampling, clustering-based sampling, and importance sampling. In FFL, sampling is adjusted to consider the data contribution and fairness of each client. It can prevent data bias or exclusion of minority opinions, and generate a more balanced model that reflects the opinions of all participants.

DivFL by Balakrishman et al. provides a thorough analysis of convergence in heterogeneous settings and shows several advantages of the approach by applying it to synthetic and real datasets, including improved learning efficiency, faster convergence, and more uniform performance across clients[106]. DELTA by Wang et al. characterizes the effects of client diversity and local dispersion and samples representative clients as information for global model updating[107]. Luo et al. designed an adaptive client sampling algorithm that addresses both systemic and statistical heterogeneity to minimize wall-clock convergence time[126].

4.3.3. Fault Tolerance

In FL, fault tolerance is used as distributed fault tolerance, redundancy techniques, and strong aggregation methods. Distributed fault tolerance creates a system that does not rely on a central server for all computations, reducing the risk of a single point of failure. Redundancy techniques implement redundant computation or data storage across nodes, allowing the system to recover from individual node failures without losing important information. The link between FL and fault tolerance is Redundancy Techniques, Checkpointing and Rollback mechanisms, and adaptive re-scheduling.

4.3.4. Model Heterogeneity

FFL provides a customized model for each client's data and computational capabilities by considering model heterogeneity. This allows all clients to optimally utilize their data to contribute to learning, and improves the fairness and performance of the overall model.

5. Fairness evaluation metrics

For the long-term viability of a fairness-aware FL system, it is critical to establish a set of performance evaluation measures. This allows us to comparatively evaluate the benefits of various proposed approaches. In general, various indicators are used in FL systems that recognize various fairness and are suitable for various scenarios and applications.

5.1. Evaluations of fairness

Fairness evaluation measures such as Average Variance, Pearson's Correlation Coefficient, and various distance measures are also used to evaluate the fairness of federated systems. If the variance of a particular system is less than the variance of another system, that system is considered fairer than the other system. Table 4 shows metrics and methods for evaluating fairness.

Table 4. Fairness evaluation metrics.

Metrics	Remarks
Accuracy	Measure how accurately the model predicts
Precision	Proportion of predicted positive outcomes that are actually positive
Recall	Proportion of true positive cases predicted by the model as positive
F1-score	Represents the harmonic average of precision and recall, and expresses the balance of the two indices
Loss Metrics	Measures how incorrectly the model predicted during training (ex: cross-entropy loss)
AUROC	Area under the receiver operating characteristic curve, measuring the classification performance of the model

MSE	Used in rare problems, the difference between the predicted and actual values is squared and averaged
MAE	An indicator that averages the absolute difference between predicted and actual values
Model Convergence Time	Measure the time it takes for the model to converge
Communication Efficiency	Measure the efficiency required for data transfer during training
Training time	Measure the time required to train a model
Model size	Size of models that need to be transferred, which affects network traffic in FFL
Update Frequency	How often the client sends model updates to the server
Resource usage	Client's CPU and memory usage during training
Scalability	Measures how well a system handles varying numbers of clients and data volumes
Data usage efficiency	Learning efficiency relative to the amount of used data
Client participation	Degree of involvement and impact of various clients
Latency	Latency between data processing and model update
Robustness	Models are resistant to data quality fluctuations or malicious attacks
Client Drift	Dispersion of trained model across clients

5.2. FFL model

Table 5 shows the state-of-the-art FFL model. It shows the latest FFL model in four categories: data-level, model level, server level, and future direction.

Table 5. Introduction to the State-Of-The-Art FFL model.

Categories	Classification	Algorithm	Remarks
Data-Level Private Data Processing	Data Preparation	Safe[127]	Detects and filters out infected data from attacked devices using clustering algorithms
		FedMix[128]	Perform data augmentation based on MixUp strategy
		Astraea[129]	Perform data augmentation based on global data distribution created by collecting local data distribution
		Faug[130]	Study the balance between personal information leakage and communication overhead through GAN-based data augmentation method
	Data Privacy Protection	PLDP-PFL[131]	Perform personalized differential privacy protection according to the sensitivity of personal data

			A Syntactic approach for privacy in FL[132]	Use anonymization techniques to reduce the sensitivity of local personal data
	External Data Utilization	Knowledge Distillation	FedMD[133]	Leverage Federated Distillation (FD) or Co-Distillation to learn knowledge from other clients
			FedGKT[134]	Through knowledge distillation, the edge knowledge of the small CNN is periodically transferred to the large server-side CNN to reduce the burden of edge learning
			FedFTG[135]	Input virtual data into global and local models for knowledge refinement
		Unsupervised Representation Learning	FedCA[136]	The FURL algorithm based on contrast loss solves the problems of data distribution inconsistency and representation inconsistency across clients
			Orchestra[137]	Discussed to learn a common representation model while decentralizing and labeling private data
			MOON[138]	Modify update direction by introducing model contrast loss
			FedProc[139]	Mitigating statistical heterogeneity through prototype-based contrastive learning
			ProtoFL[140]	Extract representations from existing models trained using existing datasets, independent of individual client data
Model-Level	Federated Optimization	Regularization	FedProx[141]	Federated optimization algorithm adding proximal flavor to FedAvg
			FedCurv[142]	Preventing serious forgetting when transferring jobs using the EWC algorithm

		pFedME[143]	Using the Moreau envelope function as a normalized loss function
	Meta Learning	Per-FedAvg[144]	A custom variant of the FedAvg algorithm based on the MAML formula
		ARUBA[145]	Leverage online convex optimization and sequence prediction algorithms to adaptively learn direct similarity and test FL performance
	Multi-task Learning	MOCHA[146]	A system-aware optimization framework for FMTL
		Ditto[147]	A scalable federated multitask learning framework with two tasks: a global goal and a local goal
Knowledge Transfer	Knowledge Distillation	FedDF[148]	Leverage unlabeled or generated data for ensemble refinement
		FedGEN[149]	Performing statistical HFL via data-free knowledge distillation method
		FedLMD[53]	Facilitates FL by recognizing different label distributions for each client
	Transfer Learning	FT-pFL[150]	Personalized knowledge transfer through knowledge coefficient matrix
		Fedhealth[151]	Federated transfer learning framework applied to the medical field
Architecture Sharing	Backbone Sharing	FedRep[152]	All clients can jointly train a global representation learning structure and then use their private data to train their own heads

		CReFF[153]	Retrain by learning the associated features, similar to training a classifier on real data
	Classifier Sharing	LG-FedAvg[154]	Extract advanced features using personalized layers and use server-shared base layers for classification
		FedPAC[155]	Reduce feature variance across clients by constraining each sample feature vector close to the global feature centroid of its category
	Other Part Sharing	HeteroFL[156]	Assign local models of various sizes depending on the computational and communication capabilities of each client
		FedLA[157]	Utilizes a hypernetwork of servers to evaluate the importance of each client model layer and generate aggregate weights for each model layer
		CD2-pFed[158]	Dynamically separate global model parameters for personalization
Server-Level	Client Selection	Favor[159]	A heuristic-based control framework that actively selects an optimal subset of clients to participate in the FL iterative process
		CUCB[160]	Client selection algorithm to minimize class imbalance and facilitate global model convergence
		FedSAE[161]	Estimate the reliability of each device and select clients based on training loss

		FedCS[162]	Perform client selection tasks based on data resources, computer capabilities, and wireless channel conditions
	Client Clustering	FL + HC[163]	Introducing a hierarchical clustering step to separate client clusters based on the similarity of client updates to the global joint model
		FeSEM[164]	Calculate the distance between local models and cluster centroids using SEM optimization
		CFL[165]	Clustering similar clients via cosine similarity between gradient updates
		FLAME[166]	Detect adversarial model updates through a clustering strategy that limits the noise scale of backdoor denoising
	Decentralized Communication	Combo[167]	After dividing the local model into model segments, randomly select some clients to send the model segments
		ProxyFL[168]	Ensures that each client maintains two models: a private model for the exchange and a publicly shared proxy model
		BFLC[169]	Strengthen the security of FL by leveraging blockchain to store global models and exchange local model updates
Future Direction	Improving Communication Efficiency	CMFL[170]	Prevents irrelevant updates from being sent to the server by measuring whether local updates are consistent with global updates
		FedDM[171]	Construct some synthetic data locally on the client to have a

		similar distribution to the original data for the loss function
	DisPFL[172]	Adopt distributed sparse learning technology
Federated Fairness	FedDM[84]	Modified method of differential multipliers
	FPFL[173]	Improving differential multiplier MMDM to improve system fairness
	q-FedAvg[92]	Improve fairness by reducing accuracy differences in client models
	CFFL[101]	Collaborative fairness is achieved in FL by assigning models with different performance according to the contribution of each client
	FFLFCN[87]	Personal information protection in the medical field FFL
	PrivFairFL[44]	The conflict between fairness and privacy is resolved by combining FL with secure multi-party computation (SMC) and differential privacy (DP)
	CVFL[174]	To alleviate the straggler problem, we design a new optimization objective that can increase the contribution of stragglers to the trained model
	Incentive design and differential privacy based federated learning[175]	Designing a new incentive mechanism to encourage many data owners to participate in the FL process through MD and DP, considering privacy protection
	FairFL[176]	It consists of a principled deep multi-agent reinforcement learning framework and a secure information aggregation protocol

	that optimizes both the accuracy and fairness of the learned model while respecting the strict privacy constraints of the client
FairVFL[21]	Learn an integrated and fair representation of samples based on distributed feature fields in a privacy-preserving manner. Specifically, each platform with fairness-insensitive features first learns a local data representation from the local features
FedFB[177]	Modifies the FedAvg protocol to effectively mimic centralized process learning
Dubhe[178]	To address the statistical heterogeneity problem, we propose a pluggable system-level client selection method called Dubhe. With the help of HE, clients actively participate in learning while protecting their personal information
FedEBA+[179]	We propose FedEBA+, a new FL algorithm that enhances fairness while improving global model performance. FedEBA+ integrates a fair aggregation system that assigns higher weights to low-performing clients and a sort update method, provides theoretical convergence analysis, and demonstrates fairness
AFLPC[180]	Reduce noise while protecting data privacy using an adaptive differential privacy mechanism. We propose a weight-based asynchronous FL aggregate update method to reasonably control the

	proportion of parameters submitted by users with different training speeds in the aggregate parameters, and actively update the aggregate parameters of delayed users to find the speed difference in the model. Effectively reduce negative impacts
Blockchain-orchestrated machine learning[181]	Exploring more detailed combinations of uses along with the auditability and incentives that blockchain can allow in machine learning processes, with a focus on decentralizing or federating the learning process. Provides an advanced blockchain-orchestrated machine learning system for privacy-preserving FL in medicine based on cost-benefit analysis and a framework for new utility in the health field
FairFed[50]	We empirically evaluate it against a common baseline for fair ML and FL, and provide a fairer model under highly heterogeneous data distributions across clients. Exploring techniques for ensuring group fairness in a FL environment and addressing potential biases that may arise during model training
FeMinMax[182]	Formally analyzing how fairness goals differ from existing FL fairness criteria that impose similar performance across participants instead of demographic groups
FGFL[183]	Evaluate players based on reputation and contributing factors

			and create a blockchain-based incentive governor for FL. Job publishers pay clients fairly to recruit efficient players, and malicious players are punished and deleted
		Bounded Group Loss[184]	We explore and extend the concept of Bounded Group Loss as a theoretically grounded approach to group fairness that offers a favorable balance over previous work between fairness and usefulness. We propose a scalable federated optimization method to optimize empirical risk under multiple group fairness constraints
		FIFL[185]	Compensate workers fairly to attract trustworthy and efficient workers and punish and eliminate malicious workers based on a dynamic, real-time worker evaluation process
Privacy Protection		DP-FedAvg[64]	Applying a Gaussian mechanism to add user-level privacy features to FedAvg
		FedMD-NFDP[186]	Adding a noise-free DP mechanism to FedMD to protect data privacy without generating noise
Attack Robustness	Attack Methods	DBA[187]	Decompose global triggers into local triggers and inject them into multiple malicious clients
		Edge-case backdoors[188]	Consider contaminating edge case samples (tail data of the data distribution)
	Defense Stratedgies	CRFL[189]	Improve robustness against backdoor attacks by clipping the model and adding soft noise

		RBML-DFL[190]	Prevent central server failure or malfunction through blockchain encrypted transactions
		ResSFL[191]	We obtain a durable feature extractor that is trained by experts with the attacker's perception to initialize the client model.
		Soteria[192]	Attack defenses are performed by creating distorted data representations, which reduces the quality of the reconstructed data.
		BaFFLe[193]	The server trains a backdoor filter and randomly sends it to the client to identify and remove backdoor instances.
Uniform Benchmark	General Federated Learning Systems	FedML[194]	A research library that supports distributed learning, mobile on-device learning, and standalone simulation learning. Provides standardized implementations of several existing FL algorithms and provides standardized benchmark settings for various datasets, including non-IID segmentation methods, number of devices, and baseline models.
		FedScale[195]	A federated learning benchmark suite that provides real-world datasets covering a wide range of FL tasks, including image classification, object detection, language modeling, and speech recognition. FedScale also includes the extensible and extensible FedScale runtime to enable and standardize real-world endpoint deployments of FL.

	OARF[63]	Simulate real-world data distribution using public datasets collected from various sources. Additionally, OARF quantitatively studies preliminary relationships between various design indicators such as data partitioning and privacy protection mechanisms in FL systems.
	FedEval[196]	FL evaluation model with five metrics including accuracy, communication, time consumption, privacy, and robustness. FedEval is implemented and evaluated on two of the most widely used algorithms: FedSGD and FedAvg.
Specific Federated Learning Systems	FedReIDBench[197]	A new benchmark for implementing FL on human ReID, including 9 different datasets and 2 federation scenarios. Specifically, the two federation scenarios are the camera federation scenario and the dataset federation scenario, which represent the standard server-client architecture and client-edge-cloud architecture, respectively.
	pFL-Bench[198]	A benchmark for personalized FL, covering 12 different dataset variants including images, text, graphs, and recommendation data with integrated data partitioning and realistic heterogeneous settings. Additionally, pFL-Bench provides implementation of over 20 competitive, personalized FL criteria to aid in standardized evaluation.

	FedGraphNN[199]	A benchmark system built on a unified formulation of graph federated learning, including extensive datasets from seven fields, popular graph neural network (GNN) models and FL algorithms.
Datasets	LEAF[200]	Contains six types of federated datasets covering a variety of fields, including image classification (FEMNIST, Synthetic Dataset), image recognition (Celeba), sentiment analysis (Sentiment140), and next character prediction (Shakespeare, Reddit). Additionally, LEAF provides two sampling methods, 'IID' and 'non-IID', to partition the dataset to different clients.
	Street Dataset[201]	Introducing a federated dataset for object detection. This dataset contains over 900 images generated from 26 street cameras and 7 object categories annotated with detailed bounding boxes. Additionally, the article provides data partitioning of 5 or 20 clients, where the data distribution is non-IID and unbalanced, reflecting the characteristics of real federated learning scenarios.

6. FFL Application in Wireless Communication

6.1. Improving Usability

FFL extends the application areas of traditional FL, especially providing an approach that enhances data fairness and privacy. Applications of FFL include usability improvements and unique use cases in wireless communications. First, in the application, FFL can improve the usability of mobile apps by learning user behavior patterns while protecting user privacy through mobile device optimization[202]. Second, in the service, it can be used in financial fraud detection, where banks and financial institutions can jointly train fraud detection models without sharing their data[203]. Third, it can be used in intelligent medical diagnosis systems, where patient data collected from various hospitals and medical institutions can be used to diagnose diseases and develop treatments[204]. By using FFL, each institution can jointly develop more accurate diagnosis models while maintaining the confidentiality of patient data. By integrating various data sources including patients' genetic information, lifestyle habits, and health conditions, it can suggest treatment plans optimized for individuals.

6.2. Wireless Communication

FFL is a technology that can efficiently train machine learning models across various network devices while maintaining data privacy and security in wireless communications. It mainly focuses on optimizing model learning from distributed data sources and minimizing network latency. In intelligent transportation systems, data collected from vehicles is used to analyze real-time traffic conditions and optimize traffic flow. In smart grid management, smart meter data is used to adjust energy demand and supply in real time. Through FFL, each smart meter processes data individually, and a central aggregator integrates them to optimize energy distribution. FFL will further expand with the advancement of wireless communication technology in the future. In particular, the increase in IoT (Internet of Things) devices and the spread of 5G networks are expected to accelerate the development of this technology. First, personalized healthcare services will provide customized healthcare services by utilizing individual patient health data. Through FFL, data can be processed locally on each patient's device and integrated to suggest personalized health advice and treatment methods. Second, in disaster response systems, data collected from various sensors and cameras can be used to build a system that can respond quickly to natural disasters. FFL processes data from each sensor locally and provides real-time information for disaster response through central aggregation.

7. FFL Challenge and Future work

7.1. Privacy Preservation

FFL further enhances privacy protection, and includes mechanisms to ensure data fairness and protect data from overfitting in minority groups. Various research and development are underway as approaches to enhance privacy protection while improving model fairness. The future direction of FFL includes innovations in technical, ethical, and legal aspects. In technical innovation, advanced encryption techniques such as HE and SMPC are currently used in FFL systems. In the future, these technologies will further develop to improve processing speed and to efficiently process more complex data types and large-scale data. Here are some approaches for this: Extended HE: Currently, homomorphic encryption has limitations in computation, but it is necessary to expand the ability to process data in an encrypted state while reducing computational costs through more advanced forms. Utilization of large-scale language models such as GPT: We seek ways to derive personalized results through pre-trained models without directly processing user data. Establishment of international data privacy standards: We can improve the level of privacy protection in FFL by establishing privacy protection standards that can be commonly recognized by various countries and companies. Strengthening transparency and accountability: Provide transparency into the decision-making process of the FFL system and clarify the criteria for holding those responsible when privacy violations occur.

7.2. Communication Cost

In order to improve communication efficiency, FFL uses an optimized algorithm to minimize the amount of data transmitted by each participant and reduce communication costs. In particular, in large-scale distributed networks, where numerous participants share data with a central server, the high communication costs incurred in the process can reduce the efficiency of the entire system. Here are some approaches for this: Model Compression and Quantization: The size of the transmitted data can be reduced by compressing or quantizing model parameters. It reduces the amount of data required for model updates, thereby reducing the communication load. Sparse Model Training: Eliminate unimportant parameters of the model or minimize updates, and only exchange practically important parameters periodically. Asynchronous FL: Participants do not perform updates simultaneously, but communicate with the server asynchronously according to their own schedules, which helps reduce network latency and communication burden. Staleness-aware Aggregation: It reduces the need for communication by effectively utilizing even old updates, and finds the optimal balance between model freshness and efficiency. Edge Computing: It is effective in processing data close to the data generation point and exchanging only essential information with the central server, thereby reducing the amount of data to be transmitted and minimizing delay time. Network Topology Optimization: By optimizing the network topology, efficient data routing and resource usage can be promoted, and by shortening the data transmission path within the network, communication costs can be reduced. Resource Allocation: By efficiently allocating communicable resources and scheduling that takes into account the communication time of participants, network congestion and communication costs can be managed.

7.3. Systems Heterogeneity

In order for participants with various hardware, software, and network environments to efficiently train models and derive fair results, it is necessary to solve the heterogeneity problem. Here are some approaches for this. Model Compression: The size and complexity of the model are adjusted according to the computing ability of the participant, and each participant can train using a model suitable for their hardware. Layer-wise Adaptation: Only some layers of the model are learned locally by the participant, and important layers are learned centrally to manage heterogeneity. Asynchronous updates: Updates are performed asynchronously according to the network speed or availability of the participant, and each participant can transmit data at their own speed, which can improve the efficiency of the entire network. Bandwidth-aware Scheduling: The data transmission schedule can be optimized by considering the bandwidth of the participant, minimizing network congestion and reducing data transmission costs. Resource-aware Task Allocation: The learning task is allocated according to the computing ability of each participant, and all participants perform tasks suitable for their devices, which contributes to improving the overall learning efficiency. On-device Machine Learning: Use lightweight machine learning models to perform learning directly on each participant's device, enabling learning even on devices with limited computing resources. Cross-platform Frameworks: Use cross-platform frameworks such as Lite or PyTorch Mobile to effectively run models on a variety of operating systems and hardware. Adaptive Code Deployment Develop tools that automatically optimize and deploy code to suit the characteristics of the platform.

7.4. Unreliable Model Upload

The problem of unstable model upload in FFL can reduce the learning efficiency and model accuracy of the entire system because participants can upload inaccurate or corrupted model updates to the central server. Here are some approaches: Pre-upload validation: Before participants upload models, a rigorous validation process is performed locally to evaluate the accuracy, consistency, and data quality of the model. Server-side validation: After receiving the uploaded model from the central server, additional validation is performed to check whether the model meets the set criteria. Quality-based incentives: Motivate participants to pay more attention to data quality and accuracy by providing greater rewards for participants who provide high-quality model updates. Penalty mechanism: Introduce penalties for participants who repeatedly submit inaccurate or corrupted updates to encourage them to provide only reliable data. Anomaly detection algorithm: Develop a machine learning-based algorithm to automatically detect outliers in uploaded model updates, so that the reliability of updates can be evaluated in an automated manner. Robust integration: When integrating updates from multiple participants, a robust integration mechanism that can tolerate outliers or errors is designed to enhance the stability of the overall model and minimize the impact of one or two inaccurate updates on the overall model. Each strategy can work complementarily to improve the overall reliability and efficiency of the system.

7.5. Multi-Center FL (MCFL)

MCFL is a method in which multiple data centers collaborate to build a common machine learning model. It is particularly difficult to secure communication, data homogeneity, and balanced data contributions between data centers. Therefore, the following approaches can be considered to effectively develop MCFL. Improved synchronization mechanism: Develop a more efficient data streaming and version management system for synchronization between data centers, and ensure that all participating centers have access to the latest data and model updates. Fair data contribution evaluation: Introduce a mechanism to fairly distribute rewards by evaluating the contribution of each data center, motivating all participating centers to contribute to the model, and improving the quality and quantity of data. Model Customization: Adopt a model structure that suits the characteristics of each data center, so that the characteristics of local data can be reflected as much as possible, which can improve the efficiency and accuracy of the overall model. Continuous model optimization: Continuously update and optimize the model through continuous feedback and performance monitoring, helping the model to quickly adapt to changing data conditions.

7.6. MMFFL Research Direction and Tasks

Data from different modalities are difficult to synchronize and can affect model learning, so it is necessary to develop a mechanism to effectively integrate features extracted from each modality. MMFFL can consider the following methods to maintain privacy while dealing with data imbalance and bias issues. Segmentation and appropriate feature extraction: Apply a feature extraction method optimized for each modality to extract the most useful information from each data type. Diversity promotion: Include data from participants with diverse backgrounds in each modality to increase the representativeness of the learning data, and ultimately promote fair model learning. Therefore, examples of future directions for MMFFL are as follows. Diagnostic systems in the medical field develop systems that provide more accurate and personalized diagnoses by integrating various forms of medical data (e.g., patient medical history information, X-ray images, and voice records). In this process, fairness of each data source is secured to provide equal treatment opportunities to all patients. Driver assistance systems in the automotive field integrate in-vehicle cameras, sensor data, and driver voice commands to learn driver behavior patterns and provide optimal driving support. This system ensures the same level of safety and convenience for all drivers by considering the fairness of each data source. It analyzes the user's text input, image upload, and interaction data through multimodal federated learning through the recommendation system of social media, and recommends personalized content. In this process, it ensures fair processing of user data, and aims to recommend unbiased content to all users. This direction will promote the development of multimodal federated learning and enable the transition to a more fair and inclusive technology. MMFFL offers the potential to develop more accurate and fair artificial intelligence models by utilizing rich and diverse data sources. However, multimodal research is still in its early stages, and future research and technological developments will further expand the possibilities of this approach.

7.7. Reliable Client Selection

FFL focuses on using only high-quality data by more thoroughly evaluating the reliability of participants, and significantly improving the fairness of data, privacy protection, and system reliability compared to existing FL. Selecting reliable clients is important for maintaining the quality of data and the accuracy and fairness of the model within the federated learning network. To solve the problems caused by the asymmetric data distribution and diverse computational capabilities among clients, the following approaches can be included. Behavior-based evaluation: Reliable clients can be identified by evaluating the client's previous participation history and the consistency of data provision. Introducing a point system: Motivate clients by awarding points whenever they continuously provide reliable data, and providing additional benefits when a certain number of points are accumulated. Feedback loop: Provide feedback on the client's data contribution so that the client source can improve the quality of data.

8. Conclusion

As the digital age exponentially increases the volume and variety of data, the need for ethical data utilization and strong privacy protection measures is growing. Fair Federated Learning (FFL) has emerged as an instrumental paradigm to address these rapidly growing needs. FFL has evolved to develop fair models that not only mitigate the risk of marginalizing individual participants or minority groups by integrating fairness into the core of its architecture, but also ensure fair benefits for all stakeholders. It also effectively addresses bias that can occur when relying on a single data source and prevents overfitting by learning from diverse data sources through bias reduction. Through this fairness enhancement and bias reduction, FFL enhances privacy by ensuring that important personal information is not transmitted over the network through local processing of data. This shift from traditional Federated Learning (FL) practices is toward correcting inherent bias and strengthening model integrity through a more comprehensive and balanced data representation.

Additionally, FFL excels at effectively deterring potential malicious interventions and data manipulation by strengthening the trustworthiness of shared models within federated networks. It advocates transparent and fair data contribution evaluation that resonates with ethical data handling practices, paving the way for safe and reliable model sharing. The evolution from FL to FFL exemplifies a holistic approach that reconciles the twin concerns of fairness and privacy, moving beyond simply optimizing model performance to embracing a governance framework that prioritizes equity and confidentiality. In essence, FFL not only advances FL methodologies, but also more closely aligns with the stringent regulatory and ethical standards that govern modern machine learning applications. This paper details the strategic implementation and challenges faced in realizing FFL, and highlights how this innovative approach can dramatically change the landscape of data-driven decision making by ensuring fairness and protecting participant privacy across a range of sectors.

Author Contributions: Conceptualization, D.K. and H.W.; methodology, D.K.; software, D.K.; validation, D.K., H.W. and Y.L.; formal analysis, D.K.; investigation, D.K.; resources, D.K.; data curation, D.K.; writing—original draft preparation, D.K.; writing—review and editing, D.K. and H.W.; visualization, D.K.; supervision, H.W. and Y.L.; project administration, D.H. and H.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(RS-2024-00350688). This work was supported by the Gachon University research fund of 2022(GCU-202300760001).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to thank the editors and anonymous reviewers for their detailed comments and efforts toward improving our research paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj* **2021**, *372*.
2. Chang, H.; Shokri, R. Bias propagation in federated learning. *arXiv preprint arXiv:2309.02160* **2023**.
3. Zawad, S.; Yan, F.; Anwar, A. Systems bias in federated learning. In *Federated Learning: A Comprehensive Overview of Methods and Applications*; Springer: 2022; pp. 259-278.
4. Rafi, T.H.; Noor, F.A.; Hussain, T.; Chae, D.-K. Fairness and privacy preserving in federated learning: A survey. *Information Fusion* **2024**, *105*, 102198.
5. Wang, H.; Xu, J. Combating client dropout in federated learning via friend model substitution. *arXiv preprint arXiv:2205.13222* **2022**.
6. Sun, Y.; Mao, Y.; Zhang, J. Mimic: Combating client dropouts in federated learning by mimicking central updates. *IEEE Transactions on Mobile Computing* **2023**.
7. Guo, Y.; Tang, X.; Lin, T. FedDebias: Reducing the Local Learning Bias Improves Federated Learning on Heterogeneous Data. **2022**.

8. Taghiyarrenani, Z.; Nowaczyk, S.; Pashami, S. Analysis of Statistical Data Heterogeneity in Federated Fault Identification. In Proceedings of the PHM Society Asia-Pacific Conference, 2023.
9. Dai, Y.; Chen, Z.; Li, J.; Heinecke, S.; Sun, L.; Xu, R. Tackling data heterogeneity in federated learning with class prototypes. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2023; pp. 7314-7322.
10. Vahidian, S.; Morafah, M.; Chen, C.; Shah, M.; Lin, B. Rethinking data heterogeneity in federated learning: Introducing a new notion and standard benchmarks. *IEEE Transactions on Artificial Intelligence* **2023**, *5*, 1386-1397.
11. Chen, L.; Zhang, W.; Dong, C.; Huang, Z.; Nie, Y.; Hou, Z.; Qiao, S.; Tan, C.W. FedDRL: Trustworthy Federated Learning Model Fusion Method Based on Staged Reinforcement Learning. *Computing and Informatics* **2024**, *43*, 1–37-31–37.
12. Qi, P.; Chiaro, D.; Guzzo, A.; Ianni, M.; Fortino, G.; Piccialli, F. Model aggregation techniques in federated learning: A comprehensive survey. *Future Generation Computer Systems* **2023**.
13. Li, T.; Sahu, A.K.; Talwalkar, A.; Smith, V. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine* **2020**, *37*, 50-60.
14. Ye, M.; Fang, X.; Du, B.; Yuen, P.C.; Tao, D. Heterogeneous federated learning: State-of-the-art and research challenges. *ACM Computing Surveys* **2023**, *56*, 1-44.
15. Zhang, J.; Wang, J.; Li, Y.; Xin, F.; Dong, F.; Luo, J.; Wu, Z. Addressing Heterogeneity in Federated Learning with Client Selection via Submodular Optimization. *ACM Transactions on Sensor Networks* **2024**, *20*, 1-32.
16. Djebrouni, Y.; Benarba, N.; Touat, O.; De Rosa, P.; Bouchenak, S.; Bonifati, A.; Felber, P.; Marangozova, V.; Schiavoni, V. Bias mitigation in federated learning for edge computing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **2024**, *7*, 1-35.
17. Ferraguig, L.; Djebrouni, Y.; Bouchenak, S.; Marangozova, V. Survey of bias mitigation in federated learning. In Proceedings of the Conférence francophone d'informatique en Parallélisme, Architecture et Système, 2021.
18. Gao, Y.; Lu, G.; Gao, J.; Li, J. A High-Performance Federated Learning Aggregation Algorithm Based on Learning Rate Adjustment and Client Sampling. *Mathematics* **2023**, *11*, 4344.
19. Han, M.; Zhu, T.; Zhou, W. Fair Federated Learning with Opposite GAN. *Knowledge-Based Systems* **2024**, *287*, 111420.
20. Zhang, H.; Zhang, P.; Hu, M.; Liu, M.; Wang, J. FedUB: Federated Learning Algorithm Based on Update Bias. *Mathematics* **2024**, *12*, 1601.
21. Qi, T.; Wu, F.; Wu, C.; Lyu, L.; Xu, T.; Liao, H.; Yang, Z.; Huang, Y.; Xie, X. Fairvfl: A fair vertical federated learning framework with contrastive adversarial learning. *Advances in neural information processing systems* **2022**, *35*, 7852-7865.
22. Liu, Y.; Kang, Y.; Zou, T.; Pu, Y.; He, Y.; Ye, X.; Ouyang, Y.; Zhang, Y.-Q.; Yang, Q. Vertical federated learning: Concepts, advances, and challenges. *IEEE Transactions on Knowledge and Data Engineering* **2024**.
23. Yang, L.; Chai, D.; Zhang, J.; Jin, Y.; Wang, L.; Liu, H.; Tian, H.; Xu, Q.; Chen, K. A survey on vertical federated learning: From a layered perspective. *arXiv preprint arXiv:2304.01829* **2023**.
24. Wu, Z.; Li, Q.; He, B. Practical vertical federated learning with unsupervised representation learning. *IEEE Transactions on Big Data* **2022**.
25. Zhang, F.; Shuai, Z.; Kuang, K.; Wu, F.; Zhuang, Y.; Xiao, J. Unified fair federated learning for digital healthcare. *Patterns* **2024**, *5*.
26. Chen, P.; Du, X.; Lu, Z.; Chai, H. Universal adversarial backdoor attacks to fool vertical federated learning. *Computers & Security* **2024**, *137*, 103601.
27. Xuan, Y.; Chen, X.; Zhao, Z.; Tang, B.; Dong, Y. Practical and general backdoor attacks against vertical federated learning. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, 2023; pp. 402-417.
28. Salazar, T.; Fernandes, M.; Araújo, H.; Abreu, P.H. Fair-fate: Fair federated learning with momentum. In Proceedings of the International Conference on Computational Science, 2023; pp. 524-538.
29. Zhou, P.; Xu, H.; Lee, L.H.; Fang, P.; Hui, P. Are you left out? an efficient and fair federated learning for personalized profiles on wearable devices of inferior networking conditions. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **2022**, *6*, 1-25.
30. Zhao, Z.; Joshi, G. A dynamic reweighting strategy for fair federated learning. In Proceedings of the ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022; pp. 8772-8776.
31. Coston, A.; Ramamurthy, K.N.; Wei, D.; Varshney, K.R.; Speakman, S.; Mustahsan, Z.; Chakraborty, S. Fair transfer learning with missing protected attributes. In Proceedings of the Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, 2019; pp. 91-98.
32. Fan, Y.; Xu, W.; Wang, H.; Zhu, J.; Guo, S. Balanced Multi-modal Federated Learning via Cross-Modal Infiltration. *arXiv preprint arXiv:2401.00894* **2023**.

33. Su, C.; Yu, G.; Wang, J.; Li, H.; Li, Q.; Yu, H. Multi-Dimensional Fair Federated Learning. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2024; pp. 15083-15090.
34. Zhou, X.; Yang, Q.; Zheng, X.; Liang, W.; Kevin, I.; Wang, K.; Ma, J.; Pan, Y.; Jin, Q. Personalized federation learning with model-contrastive learning for multi-modal user modeling in human-centric metaverse. *IEEE Journal on Selected Areas in Communications* **2024**.
35. Paillier, P. Public-key cryptosystems based on composite degree residuosity classes. In Proceedings of the International conference on the theory and applications of cryptographic techniques, 1999; pp. 223-238.
36. Fan, J.; Vercauteren, F. Somewhat practical fully homomorphic encryption. *Cryptology ePrint Archive* **2012**.
37. Jiang, Z.; Wang, W.; Liu, Y. Flashe: Additively symmetric homomorphic encryption for cross-silo federated learning. *arXiv preprint arXiv:2109.00675* **2021**.
38. Jin, W.; Yao, Y.; Han, S.; Joe-Wong, C.; Ravi, S.; Avestimehr, S.; He, C. FedML-HE: An efficient homomorphic-encryption-based privacy-preserving federated learning system. *arXiv preprint arXiv:2303.10837* **2023**.
39. Volgushev, N.; Schwarzkopf, M.; Getchell, B.; Varia, M.; Lapets, A.; Bestavros, A. Conclave: secure multi-party computation on big data. In Proceedings of the Proceedings of the Fourteenth EuroSys Conference 2019, 2019; pp. 1-18.
40. Li, Y.; Zhou, Y.; Jolfaei, A.; Yu, D.; Xu, G.; Zheng, X. Privacy-preserving federated learning framework based on chained secure multiparty computing. *IEEE Internet of Things Journal* **2020**, *8*, 6178-6186.
41. Kalapaaking, A.P.; Stephanie, V.; Khalil, I.; Atiquzzaman, M.; Yi, X.; Almashor, M. Smpc-based federated learning for 6g-enabled internet of medical things. *IEEE Network* **2022**, *36*, 182-189.
42. Li, J.; Yan, T.; Ren, P. VFL-R: a novel framework for multi-party in vertical federated learning. *Applied Intelligence* **2023**, *53*, 12399-12415.
43. Mohassel, P.; Zhang, Y. Secureml: A system for scalable privacy-preserving machine learning. In Proceedings of the 2017 IEEE symposium on security and privacy (SP), 2017; pp. 19-38.
44. Pentyala, S.; Neophytou, N.; Nascimento, A.; De Cock, M.; Farnadi, G. Privfairfl: Privacy-preserving group fairness in federated learning. *arXiv preprint arXiv:2205.11584* **2022**.
45. Mugunthan, V.; Polychroniadou, A.; Byrd, D.; Balch, T.H. Smpai: Secure multi-party computation for federated learning. In Proceedings of the Proceedings of the NeurIPS 2019 Workshop on Robust AI in Financial Services, 2019.
46. Chen, L.; Xiao, D.; Yu, Z.; Zhang, M. Secure and efficient federated learning via novel multi-party computation and compressed sensing. *Information Sciences* **2024**, *667*, 120481.
47. Gao, D.; Liu, Y.; Huang, A.; Ju, C.; Yu, H.; Yang, Q. Privacy-preserving heterogeneous federated transfer learning. In Proceedings of the 2019 IEEE international conference on big data (Big Data), 2019; pp. 2552-2559.
48. Liu, Y.; Ma, Z.; Yan, Z.; Wang, Z.; Liu, X.; Ma, J. Privacy-preserving federated k-means for proactive caching in next generation cellular networks. *Information Sciences* **2020**, *521*, 14-31.
49. Xu, G.; Li, H.; Liu, S.; Yang, K.; Lin, X. VerifyNet: Secure and verifiable federated learning. *IEEE Transactions on Information Forensics and Security* **2019**, *15*, 911-926.
50. Ezzeldin, Y.H.; Yan, S.; He, C.; Ferrara, E.; Avestimehr, A.S. Fairfed: Enabling group fairness in federated learning. In Proceedings of the Proceedings of the AAAI conference on artificial intelligence, 2023; pp. 7494-7502.
51. Li, A.; Sun, J.; Zeng, X.; Zhang, M.; Li, H.; Chen, Y. Fedmask: Joint computation and communication-efficient personalized federated learning via heterogeneous masking. In Proceedings of the Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems, 2021; pp. 42-55.
52. Setayesh, M.; Li, X.; Wong, V.W. Perfedmask: Personalized federated learning with optimized masking vectors. In Proceedings of the The Eleventh International Conference on Learning Representations, 2023.
53. Lu, J.; Li, S.; Bao, K.; Wang, P.; Qian, Z.; Ge, S. Federated Learning with Label-Masking Distillation. In Proceedings of the Proceedings of the 31st ACM International Conference on Multimedia, 2023; pp. 222-232.
54. He, S.; Yan, Q.; Wu, F.; Wang, L.; Lécuyer, M.; Beschastnikh, I. Gluefl: Reconciling client sampling and model masking for bandwidth efficient federated learning. *Proceedings of Machine Learning and Systems* **2023**, *5*, 695-707.
55. Xiong, J.; Zhu, H. PrivMaskFL: A private masking approach for heterogeneous federated learning in IoT. *Computer Communications* **2024**, *214*, 100-112.
56. Girgis, A.; Data, D.; Diggavi, S.; Kairouz, P.; Suresh, A.T. Shuffled model of differential privacy in federated learning. In Proceedings of the International Conference on Artificial Intelligence and Statistics, 2021; pp. 2521-2529.
57. Liu, R.; Cao, Y.; Chen, H.; Guo, R.; Yoshikawa, M. Flame: Differentially private federated learning in the shuffle model. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2021; pp. 8688-8696.

58. Zhou, Z.; Xu, C.; Wang, M.; Kuang, X.; Zhuang, Y.; Yu, S. A multi-shuffler framework to establish mutual confidence for secure federated learning. *IEEE Transactions on Dependable and Secure Computing* **2022**, *20*, 4230-4244.
59. Liu, Y.; Wang, Z.; Zhu, Y.; Chen, C. DPBalance: Efficient and Fair Privacy Budget Scheduling for Federated Learning as a Service. *arXiv preprint arXiv:2402.09715* **2024**.
60. ur Rehman, M.H.; Dirir, A.M.; Salah, K.; Svetinovic, D. Fairfed: Cross-device fair federated learning. In Proceedings of the 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), 2020; pp. 1-7.
61. Bhowmick, A.; Duchi, J.; Freudiger, J.; Kapoor, G.; Rogers, R. Protection against reconstruction and its applications in private federated learning. *arXiv preprint arXiv:1812.00984* **2018**.
62. Zhang, J.; Chen, B.; Yu, S.; Deng, H. PEFL: A privacy-enhanced federated learning scheme for big data analytics. In Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), 2019; pp. 1-6.
63. Hu, S.; Li, Y.; Liu, X.; Li, Q.; Wu, Z.; He, B. The oarf benchmark suite: Characterization and implications for federated learning systems. *ACM Transactions on Intelligent Systems and Technology (TIST)* **2022**, *13*, 1-32.
64. McMahan, H.B.; Ramage, D.; Talwar, K.; Zhang, L. Learning differentially private recurrent language models. *arXiv preprint arXiv:1710.06963* **2017**.
65. Geyer, R.C.; Klein, T.; Nabi, M. Differentially private federated learning: A client level perspective. *arXiv preprint arXiv:1712.07557* **2017**.
66. Truex, S.; Baracaldo, N.; Anwar, A.; Steinke, T.; Ludwig, H.; Zhang, R.; Zhou, Y. A hybrid approach to privacy-preserving federated learning. In Proceedings of the Proceedings of the 12th ACM workshop on artificial intelligence and security, 2019; pp. 1-11.
67. Ling, X.; Fu, J.; Chen, Z.; Wang, K.; Li, H.; Cheng, T.; Xu, G.; Li, Q. FedFDP: Federated Learning with Fairness and Differential Privacy. *arXiv preprint arXiv:2402.16028* **2024**.
68. Salim, S.; Moustafa, N.; Turnbull, B.; Razzak, I. Perturbation-enabled deep federated learning for preserving internet of things-based social networks. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **2022**, *18*, 1-19.
69. Wu, N.; Kuang, Z.; Yan, Z.; Yu, L. From Optimization to Generalization: Fair Federated Learning against Quality Shift via Inter-Client Sharpness Matching. *arXiv preprint arXiv:2404.17805* **2024**.
70. Tang, X.; Shen, M.; Li, Q.; Zhu, L.; Xue, T.; Qu, Q. Pile: Robust privacy-preserving federated learning via verifiable perturbations. *IEEE Transactions on Dependable and Secure Computing* **2023**, *20*, 5005-5023.
71. Hu, R.; Guo, Y.; Gong, Y. Federated learning with sparsified model perturbation: Improving accuracy under client-level differential privacy. *IEEE Transactions on Mobile Computing* **2023**.
72. Li, Z.; Chen, H.; Gao, Y.; Ni, Z.; Xue, H.; Shao, H. Staged noise perturbation for privacy-preserving federated learning. *IEEE Transactions on Sustainable Computing* **2024**.
73. Chamikara, M.A.P.; Bertok, P.; Khalil, I.; Liu, D.; Camtepe, S. Privacy preserving distributed machine learning with federated learning. *Computer Communications* **2021**, *171*, 112-125.
74. Zhang, Y.; Wang, Z.; Cao, J.; Hou, R.; Meng, D. ShuffleFL: Gradient-preserving federated learning using trusted execution environment. In Proceedings of the Proceedings of the 18th ACM international conference on computing frontiers, 2021; pp. 161-168.
75. Zhang, J.; Cheng, X.; Wang, W.; Yang, L.; Hu, J.; Chen, K. {FLASH}: Towards a high-performance hardware acceleration architecture for cross-silo federated learning. In Proceedings of the 20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23), 2023; pp. 1057-1079.
76. Mondal, A.; More, Y.; Rooparaghunath, R.H.; Gupta, D. Poster: Flatee: Federated learning across trusted execution environments. In Proceedings of the 2021 IEEE European Symposium on Security and Privacy (EuroS&P), 2021; pp. 707-709.
77. Ye, Y.; Li, S.; Liu, F.; Tang, Y.; Hu, W. EdgeFed: Optimized federated learning based on edge computing. *IEEE Access* **2020**, *8*, 209191-209198.
78. Issa, W.; Moustafa, N.; Turnbull, B.; Choo, K.-K.R. RVE-PFL: Robust Variational Encoder-based Personalised Federated Learning against Model Inversion Attacks. *IEEE Transactions on Information Forensics and Security* **2024**.
79. Xu, R.; Baracaldo, N.; Zhou, Y.; Anwar, A.; Ludwig, H. Hybridalpha: An efficient approach for privacy-preserving federated learning. In Proceedings of the Proceedings of the 12th ACM workshop on artificial intelligence and security, 2019; pp. 13-23.
80. Ma, J.; Naas, S.A.; Sigg, S.; Lyu, X. Privacy-preserving federated learning based on multi-key homomorphic encryption. *International Journal of Intelligent Systems* **2022**, *37*, 5880-5901.
81. Yazdinejad, A.; Dehghantanha, A.; Srivastava, G.; Karimipour, H.; Parizi, R.M. Hybrid privacy preserving federated learning against irregular users in next-generation Internet of Things. *Journal of Systems Architecture* **2024**, *148*, 103088.
82. Yao, R.; Tang, K.; Fan, B. APPFed: A Hybrid Privacy-Preserving Framework for Federated Learning over Sensitive Data. In Proceedings of the 2022 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), 2022; pp. 389-395.

83. Sánchez, P.M.S.; Celdrán, A.H.; Xie, N.; Bovet, G.; Pérez, G.M.; Stiller, B. Federatedtrust: A solution for trustworthy federated learning. *Future Generation Computer Systems* **2024**, *152*, 83-98.
84. Rodríguez-Gálvez, B.; Granqvist, F.; van Dalen, R.; Seigel, M. Enforcing fairness in private federated learning via the modified method of differential multipliers. *arXiv preprint arXiv:2109.08604* **2021**.
85. 김도형; 오경수; 이영호. HFAD: 공정한 연합학습 및 하이브리드 융합 멀티모달 산업 이상 탐지. *한국정보통신학회논문지* **2024**, *28*, 805-814.
86. Papernot, N.; Abadi, M.; Erlingsson, U.; Goodfellow, I.; Talwar, K. Semi-supervised knowledge transfer for deep learning from private training data. *arXiv preprint arXiv:1610.05755* **2016**.
87. Kim, D.-h.; Oh, K.; Kang, S.-h.; Lee, Y. Development of Pneumonia Patient Classification Model Using Fair Federated Learning. In Proceedings of the International Conference on Intelligent Human Computer Interaction, 2023; pp. 153-164.
88. Zhao, X.; Shen, D. FedSW: Federated learning with adaptive sample weights. *Information Sciences* **2024**, *654*, 119873.
89. Cellamare, M.; van Gestel, A.J.; Alradhi, H.; Martin, F.; Moncada-Torres, A. A federated generalized linear model for privacy-preserving analysis. *Algorithms* **2022**, *15*, 243.
90. Ibrahim Khalaf, O.; Algburi, S.; Selvaraj, D.; Sharif, M.S.; Elmedany, W. Federated learning with hybrid differential privacy for secure and reliable cross-IoT platform knowledge sharing. *Security and Privacy* **2024**, *7*, e374.
91. Yu, S.; Muñoz, J.P.; Jannesari, A. Federated foundation models: Privacy-preserving and collaborative learning for large models. *arXiv preprint arXiv:2305.11414* **2023**.
92. Li, T.; Sanjabi, M.; Beirami, A.; Smith, V. Fair resource allocation in federated learning. *arXiv preprint arXiv:1905.10497* **2019**.
93. Ma, J.; Tu, A.; Chen, Y.; Reddi, V.J. FedStaleWeight: Buffered Asynchronous Federated Learning with Fair Aggregation via Staleness Reweighting. *arXiv preprint arXiv:2406.02877* **2024**.
94. Yin, Q.; Huang, J.; Yao, H.; Zhang, L. Distribution-Free Fair Federated Learning with Small Samples. *arXiv preprint arXiv:2402.16158* **2024**.
95. Li, Z.; Zhou, Y.; Wu, D.; Tang, T.; Wang, R. Fairness-aware federated learning with unreliable links in resource-constrained Internet of things. *IEEE Internet of Things Journal* **2022**, *9*, 17359-17371.
96. He, C.; Ceyani, E.; Balasubramanian, K.; Annavaram, M.; Avestimehr, S. Spreadgnn: Decentralized multi-task federated learning for graph neural networks on molecular data. In Proceedings of the Proceedings of the AAAI conference on artificial intelligence, 2022; pp. 6865-6873.
97. Sprague, M.R.; Jalalirad, A.; Scavuzzo, M.; Capota, C.; Neun, M.; Do, L.; Kopp, M. Asynchronous federated learning for geospatial applications. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, 2018; pp. 21-28.
98. Nguyen, J.; Malik, K.; Zhan, H.; Yousefpour, A.; Rabbat, M.; Malek, M.; Huba, D. Federated learning with buffered asynchronous aggregation. In Proceedings of the International Conference on Artificial Intelligence and Statistics, 2022; pp. 3581-3607.
99. Chen, Y.; Ning, Y.; Slawski, M.; Rangwala, H. Asynchronous online federated learning for edge devices with non-iid data. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), 2020; pp. 15-24.
100. Baccarelli, E.; Scarpiniti, M.; Momenzadeh, A.; Ahrabi, S.S. AFAFed—Asynchronous fair adaptive federated learning for IoT stream applications. *Computer Communications* **2022**, *195*, 376-402.
101. Lyu, L.; Xu, X.; Wang, Q.; Yu, H. Collaborative fairness in federated learning. *Federated Learning: Privacy and Incentive* **2020**, 189-204.
102. Yu, H.; Liu, Z.; Liu, Y.; Chen, T.; Cong, M.; Weng, X.; Niyato, D.; Yang, Q. A fairness-aware incentive scheme for federated learning. In Proceedings of the Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 2020; pp. 393-399.
103. Bey, R.; Goussault, R.; Grolleau, F.; Benchoufi, M.; Porcher, R. Fold-stratified cross-validation for unbiased and privacy-preserving federated learning. *Journal of the American Medical Informatics Association* **2020**, *27*, 1244-1251.
104. Lu, C.; Ma, W.; Wang, R.; Deng, S.; Wu, Y. Federated learning based on stratified sampling and regularization. *Complex & Intelligent Systems* **2023**, *9*, 2081-2099.
105. Zhu, Z.; Shi, Y.; Fan, P.; Peng, C.; Letaief, K.B. ISFL: Federated Learning for Non-iid Data with Local Importance Sampling. *IEEE Internet of Things Journal* **2024**.
106. Balakrishnan, R.; Li, T.; Zhou, T.; Himayat, N.; Smith, V.; Bilmes, J. Diverse client selection for federated learning via submodular maximization. In Proceedings of the International Conference on Learning Representations, 2022.
107. Wang, L.; Guo, Y.; Lin, T.; Tang, X. Delta: Diverse client sampling for fasting federated learning. *Advances in Neural Information Processing Systems* **2024**, *36*.
108. Morell, J.Á.; Alba, E. Dynamic and adaptive fault-tolerant asynchronous federated learning using volunteer edge devices. *Future Generation Computer Systems* **2022**, *133*, 53-67.

109. Chen, J.-H.; Chen, M.-R.; Zeng, G.-Q.; Weng, J.-S. BDFL: A byzantine-fault-tolerance decentralized federated learning method for autonomous vehicle. *IEEE Transactions on Vehicular Technology* **2021**, *70*, 8639-8652.
110. Huang, W.; Ye, M.; Du, B. Learn from others and be yourself in heterogeneous federated learning. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022; pp. 10143-10153.
111. Mendieta, M.; Yang, T.; Wang, P.; Lee, M.; Ding, Z.; Chen, C. Local learning matters: Rethinking data heterogeneity in federated learning. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022; pp. 8397-8406.
112. Alam, S.; Liu, L.; Yan, M.; Zhang, M. Fedrolex: Model-heterogeneous federated learning with rolling sub-model extraction. *Advances in neural information processing systems* **2022**, *35*, 29677-29690.
113. Wang, G.; Payani, A.; Lee, M.; Kompella, R. Mitigating group bias in federated learning: Beyond local fairness. *arXiv preprint arXiv:2305.09931* **2023**.
114. Li, J.; Zhu, T.; Ren, W.; Raymond, K.-K. Improve individual fairness in federated learning via adversarial training. *Computers & Security* **2023**, *132*, 103336.
115. Commission, P.I.P. Pseudonymization information processing guidelines. **2024**, 241.
116. Welfare, M.o.H.a.; Commission, P.I.P. Healthcare data utilization guideline. **2024**, 126.
117. Turgay, S.; İlter, İ. Perturbation methods for protecting data privacy: A review of techniques and applications. *Automation and Machine Learning* **2023**, *4*, 31-41.
118. Cho, S.-G.; Lee, M.-U.; Lee, T.-H. Efficient Robust Design Optimization Using Statistical Moments Based on Multiplicative Decomposition Method. *Transactions of the Korean Society of Mechanical Engineers A* **2012**, *36*, 1109-1114.
119. Guo, J.; Pietzuch, P.; Paverd, A.; Vaswani, K. Trustworthy AI Using Confidential Federated Learning. *Communications of the ACM*.
120. Tran, A.T.; Luong, T.D.; Pham, X.S. A Novel Privacy-Preserving Federated Learning Model Based on Secure Multi-party Computation. In Proceedings of the International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making, 2023; pp. 321-333.
121. Chatzikonstantinou, C.; Konstantinidis, D.; Dimitropoulos, K.; Daras, P. Federated Learning Aggregation based on Weight Distribution Analysis. In Proceedings of the 2023 IEEE International Conference on Imaging Systems and Techniques (IST), 2023; pp. 1-6.
122. Dilley, O.; Parra-Ullauri, J.M.; Hussain, R.; Simeonidou, D. Federated Fairness Analytics: Quantifying Fairness in Federated Learning. *arXiv preprint arXiv:2408.08214* **2024**.
123. Chen, Y.; Huang, W.; Ye, M. Fair Federated Learning under Domain Skew with Local Consistency and Domain Diversity. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024; pp. 12077-12086.
124. Wang, Z.; Zhang, Z.; Tian, Y.; Yang, Q.; Shan, H.; Wang, W.; Quek, T.Q. Asynchronous federated learning over wireless communication networks. *IEEE Transactions on Wireless Communications* **2022**, *21*, 6961-6978.
125. Chen, M.; Mao, B.; Ma, T. Efficient and robust asynchronous federated learning with stragglers. In Proceedings of the International Conference on Learning Representations, 2019.
126. Luo, B.; Xiao, W.; Wang, S.; Huang, J.; Tassiulas, L. Tackling system and statistical heterogeneity for federated learning with adaptive client sampling. In Proceedings of the IEEE INFOCOM 2022-IEEE conference on computer communications, 2022; pp. 1739-1748.
127. Xu, X.; Li, H.; Li, Z.; Zhou, X. Safe: Synergic data filtering for federated learning in cloud-edge computing. *IEEE Transactions on Industrial Informatics* **2022**, *19*, 1655-1665.
128. Yoon, T.; Shin, S.; Hwang, S.J.; Yang, E. Fedmix: Approximation of mixup under mean augmented federated learning. *arXiv preprint arXiv:2107.00233* **2021**.
129. Duan, M.; Liu, D.; Chen, X.; Tan, Y.; Ren, J.; Qiao, L.; Liang, L. Astraea: Self-balancing federated learning for improving classification accuracy of mobile deep learning applications. In Proceedings of the 2019 IEEE 37th international conference on computer design (ICCD), 2019; pp. 246-254.
130. Jeong, E.; Oh, S.; Kim, H.; Park, J.; Bennis, M.; Kim, S.-L. Communication-efficient on-device machine learning: Federated distillation and augmentation under non-iid private data. *arXiv preprint arXiv:1811.11479* **2018**.
131. Hu, R.; Guo, Y.; Li, H.; Pei, Q.; Gong, Y. Personalized federated learning with differential privacy. *IEEE Internet of Things Journal* **2020**, *7*, 9530-9539.
132. Choudhury, O.; Gkoulalas-Divanis, A.; Salonidis, T.; Sylla, I.; Park, Y.; Hsu, G.; Das, A. Anonymizing data for privacy-preserving federated learning. *arXiv preprint arXiv:2002.09096* **2020**.
133. Li, D.; Wang, J. Fedmd: Heterogenous federated learning via model distillation. *arXiv preprint arXiv:1910.03581* **2019**.
134. He, C.; Annavaram, M.; Avestimehr, S. Group knowledge transfer: Federated learning of large cnns at the edge. *Advances in Neural Information Processing Systems* **2020**, *33*, 14068-14080.

135. Zhang, L.; Shen, L.; Ding, L.; Tao, D.; Duan, L.-Y. Fine-tuning global model via data-free knowledge distillation for non-iid federated learning. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022; pp. 10174-10183.
136. Zhang, F.; Kuang, K.; Chen, L.; You, Z.; Shen, T.; Xiao, J.; Zhang, Y.; Wu, C.; Wu, F.; Zhuang, Y. Federated unsupervised representation learning. *Frontiers of Information Technology & Electronic Engineering* **2023**, *24*, 1181-1193.
137. Lubana, E.S.; Tang, C.I.; Kawsar, F.; Dick, R.P.; Mathur, A. Orchestra: Unsupervised federated learning via globally consistent clustering. *arXiv preprint arXiv:2205.11506* **2022**.
138. Li, Q.; He, B.; Song, D. Model-contrastive federated learning. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021; pp. 10713-10722.
139. Mu, X.; Shen, Y.; Cheng, K.; Geng, X.; Fu, J.; Zhang, T.; Zhang, Z. Fedproc: Prototypical contrastive federated learning on non-iid data. *Future Generation Computer Systems* **2023**, *143*, 93-104.
140. Kim, H.; Kwak, Y.; Jung, M.; Shin, J.; Kim, Y.; Kim, C. Protofl: Unsupervised federated learning via prototypical distillation. In Proceedings of the Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023; pp. 6470-6479.
141. Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems* **2020**, *2*, 429-450.
142. Shoham, N.; Avidor, T.; Keren, A.; Israel, N.; Benditkis, D.; Mor-Yosef, L.; Zeitak, I. Overcoming forgetting in federated learning on non-iid data. *arXiv preprint arXiv:1910.07796* **2019**.
143. T Dinh, C.; Tran, N.; Nguyen, J. Personalized federated learning with moreau envelopes. *Advances in neural information processing systems* **2020**, *33*, 21394-21405.
144. 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.
145. Khodak, M.; Balcan, M.-F.F.; Talwalkar, A.S. Adaptive gradient-based meta-learning methods. *Advances in Neural Information Processing Systems* **2019**, *32*.
146. Smith, V.; Chiang, C.-K.; Sanjabi, M.; Talwalkar, A.S. Federated multi-task learning. *Advances in neural information processing systems* **2017**, *30*.
147. Li, T.; Hu, S.; Beirami, A.; Smith, V. Ditto: Fair and robust federated learning through personalization. In Proceedings of the International conference on machine learning, 2021; pp. 6357-6368.
148. Lin, T.; Kong, L.; Stich, S.U.; Jaggi, M. Ensemble distillation for robust model fusion in federated learning. *Advances in neural information processing systems* **2020**, *33*, 2351-2363.
149. Zhu, Z.; Hong, J.; Zhou, J. Data-free knowledge distillation for heterogeneous federated learning. In Proceedings of the International conference on machine learning, 2021; pp. 12878-12889.
150. Zhang, J.; Guo, S.; Ma, X.; Wang, H.; Xu, W.; Wu, F. Parameterized knowledge transfer for personalized federated learning. *Advances in Neural Information Processing Systems* **2021**, *34*, 10092-10104.
151. Chen, Y.; Qin, X.; Wang, J.; Yu, C.; Gao, W. Fedhealth: A federated transfer learning framework for wearable healthcare. *IEEE Intelligent Systems* **2020**, *35*, 83-93.
152. Collins, L.; Hassani, H.; Mokhtari, A.; Shakkottai, S. Exploiting shared representations for personalized federated learning. In Proceedings of the International conference on machine learning, 2021; pp. 2089-2099.
153. Shang, X.; Lu, Y.; Huang, G.; Wang, H. Federated learning on heterogeneous and long-tailed data via classifier re-training with federated features. *arXiv preprint arXiv:2204.13399* **2022**.
154. Liang, P.P.; Liu, T.; Ziyin, L.; Allen, N.B.; Auerbach, R.P.; Brent, D.; Salakhutdinov, R.; Morency, L.-P. Think locally, act globally: Federated learning with local and global representations. *arXiv preprint arXiv:2001.01523* **2020**.
155. Xu, J.; Tong, X.; Huang, S.-L. Personalized federated learning with feature alignment and classifier collaboration. *arXiv preprint arXiv:2306.11867* **2023**.
156. Diao, E.; Ding, J.; Tarokh, V. Heterofl: Computation and communication efficient federated learning for heterogeneous clients. *arXiv preprint arXiv:2010.01264* **2020**.
157. Ma, X.; Zhang, J.; Guo, S.; Xu, W. Layer-wised model aggregation for personalized federated learning. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022; pp. 10092-10101.
158. Shen, Y.; Zhou, Y.; Yu, L. Cd2-pfed: Cyclic distillation-guided channel decoupling for model personalization in federated learning. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022; pp. 10041-10050.
159. Wang, H.; Kaplan, Z.; Niu, D.; Li, B. Optimizing federated learning on non-iid data with reinforcement learning. In Proceedings of the IEEE INFOCOM 2020-IEEE conference on computer communications, 2020; pp. 1698-1707.
160. Yang, M.; Wang, X.; Zhu, H.; Wang, H.; Qian, H. Federated learning with class imbalance reduction. In Proceedings of the 2021 29th European Signal Processing Conference (EUSIPCO), 2021; pp. 2174-2178.

161. Li, L.; Duan, M.; Liu, D.; Zhang, Y.; Ren, A.; Chen, X.; Tan, Y.; Wang, C. FedSAE: A novel self-adaptive federated learning framework in heterogeneous systems. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), 2021; pp. 1-10.
162. Nishio, T.; Yonetani, R. Client selection for federated learning with heterogeneous resources in mobile edge. In Proceedings of the ICC 2019-2019 IEEE international conference on communications (ICC), 2019; pp. 1-7.
163. Briggs, C.; Fan, Z.; Andras, P. Federated learning with hierarchical clustering of local updates to improve training on non-IID data. In Proceedings of the 2020 international joint conference on neural networks (IJCNN), 2020; pp. 1-9.
164. Long, G.; Xie, M.; Shen, T.; Zhou, T.; Wang, X.; Jiang, J. Multi-center federated learning: clients clustering for better personalization. *World Wide Web* **2023**, *26*, 481-500.
165. Sattler, F.; Müller, K.-R.; Samek, W. Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints. *IEEE transactions on neural networks and learning systems* **2020**, *32*, 3710-3722.
166. Nguyen, T.D.; Rieger, P.; De Viti, R.; Chen, H.; Brandenburg, B.B.; Yalame, H.; Möllering, H.; Fereidooni, H.; Marchal, S.; Miettinen, M. {FLAME}: Taming backdoors in federated learning. In Proceedings of the 31st USENIX Security Symposium (USENIX Security 22), 2022; pp. 1415-1432.
167. Hu, C.; Jiang, J.; Wang, Z. Decentralized federated learning: A segmented gossip approach. *arXiv preprint arXiv:1908.07782* **2019**.
168. Kalra, S.; Wen, J.; Cresswell, J.C.; Volkovs, M.; Tizhoosh, H.R. Decentralized federated learning through proxy model sharing. *Nature communications* **2023**, *14*, 2899.
169. Li, Y.; Chen, C.; Liu, N.; Huang, H.; Zheng, Z.; Yan, Q. A blockchain-based decentralized federated learning framework with committee consensus. *IEEE Network* **2020**, *35*, 234-241.
170. Luping, W.; Wei, W.; Bo, L. CMFL: Mitigating communication overhead for federated learning. In Proceedings of the 2019 IEEE 39th international conference on distributed computing systems (ICDCS), 2019; pp. 954-964.
171. Xiong, Y.; Wang, R.; Cheng, M.; Yu, F.; Hsieh, C.-J. Feddm: Iterative distribution matching for communication-efficient federated learning. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023; pp. 16323-16332.
172. Dai, R.; Shen, L.; He, F.; Tian, X.; Tao, D. Dispfl: Towards communication-efficient personalized federated learning via decentralized sparse training. In Proceedings of the International conference on machine learning, 2022; pp. 4587-4604.
173. Galli, F.; Biswas, S.; Jung, K.; Cucinotta, T.; Palamidessi, C. Group privacy for personalized federated learning. *arXiv preprint arXiv:2206.03396* **2022**.
174. Xia, W.; Li, Y.; Zhang, L.; Wu, Z.; Yuan, X. A vertical federated learning framework for horizontally partitioned labels. *arXiv preprint arXiv:2106.10056* **2021**.
175. Kim, S. Incentive design and differential privacy based federated learning: A mechanism design perspective. *IEEE Access* **2020**, *8*, 187317-187325.
176. Zhang, D.Y.; Kou, Z.; Wang, D. Fairfl: A fair federated learning approach to reducing demographic bias in privacy-sensitive classification models. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), 2020; pp. 1051-1060.
177. Zeng, Y.; Chen, H.; Lee, K. Improving fairness via federated learning. *arXiv preprint arXiv:2110.15545* **2021**.
178. Zhang, S.; Li, Z.; Chen, Q.; Zheng, W.; Leng, J.; Guo, M. Dubhe: Towards data unbiasedness with homomorphic encryption in federated learning client selection. In Proceedings of the Proceedings of the 50th International Conference on Parallel Processing, 2021; pp. 1-10.
179. Wang, L.; Wang, Z.; Tang, X. Fedeba+: Towards fair and effective federated learning via entropy-based model. *arXiv preprint arXiv:2301.12407* **2023**.
180. Huang, J.; Xu, C.; Ji, Z.; Xiao, S.; Liu, T.; Ma, N.; Zhou, Q. [Retracted] AFLPC: An Asynchronous Federated Learning Privacy-Preserving Computing Model Applied to 5G-V2X. *Security and Communication Networks* **2022**, *2022*, 9334943.
181. Passerat-Palmbach, J.; Farnan, T.; McCoy, M.; Harris, J.D.; Manion, S.T.; Flannery, H.L.; Gleim, B. Blockchain-orchestrated machine learning for privacy preserving federated learning in electronic health data. In Proceedings of the 2020 IEEE international conference on blockchain (Blockchain), 2020; pp. 550-555.
182. Papadaki, A.; Martinez, N.; Bertran, M.; Sapiro, G.; Rodrigues, M. Federating for learning group fair models. *arXiv preprint arXiv:2110.01999* **2021**.
183. Li, X.; Zhao, S.; Chen, C.; Zheng, Z. Heterogeneity-aware fair federated learning. *Information Sciences* **2023**, *619*, 968-986.
184. Hu, S.; Wu, Z.S.; Smith, V. Fair federated learning via bounded group loss. In Proceedings of the 2024 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML), 2024; pp. 140-160.

185. Gao, L.; Li, L.; Chen, Y.; Zheng, W.; Xu, C.; Xu, M. Fifl: A fair incentive mechanism for federated learning. In Proceedings of the Proceedings of the 50th International Conference on Parallel Processing, 2021; pp. 1-10.
186. Sun, L.; Lyu, L. Federated model distillation with noise-free differential privacy. *arXiv preprint arXiv:2009.05537* **2020**.
187. Xie, C.; Huang, K.; Chen, P.-Y.; Li, B. Dba: Distributed backdoor attacks against federated learning. In Proceedings of the International conference on learning representations, 2019.
188. Wang, H.; Sreenivasan, K.; Rajput, S.; Vishwakarma, H.; Agarwal, S.; Sohn, J.-y.; Lee, K.; Papailiopoulos, D. Attack of the tails: Yes, you really can backdoor federated learning. *Advances in Neural Information Processing Systems* **2020**, *33*, 16070-16084.
189. Xie, C.; Chen, M.; Chen, P.-Y.; Li, B. Crfl: Certifiably robust federated learning against backdoor attacks. In Proceedings of the International Conference on Machine Learning, 2021; pp. 11372-11382.
190. Wu, D.; Wang, N.; Zhang, J.; Zhang, Y.; Xiang, Y.; Gao, L. A blockchain-based multi-layer decentralized framework for robust federated learning. In Proceedings of the 2022 international joint conference on neural networks (IJCNN), 2022; pp. 1-8.
191. Li, J.; Rakin, A.S.; Chen, X.; He, Z.; Fan, D.; Chakrabarti, C. Ressfl: A resistance transfer framework for defending model inversion attack in split federated learning. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022; pp. 10194-10202.
192. Sun, J.; Li, A.; Wang, B.; Yang, H.; Li, H.; Chen, Y. Soteria: Provable defense against privacy leakage in federated learning from representation perspective. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021; pp. 9311-9319.
193. Andreina, S.; Marson, G.A.; Möllering, H.; Karame, G. Baffle: Backdoor detection via feedback-based federated learning. In Proceedings of the 2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS), 2021; pp. 852-863.
194. He, C.; Li, S.; So, J.; Zeng, X.; Zhang, M.; Wang, H.; Wang, X.; Vepakomma, P.; Singh, A.; Qiu, H. Fedml: A research library and benchmark for federated machine learning. *arXiv preprint arXiv:2007.13518* **2020**.
195. Lai, F.; Dai, Y.; Singapuram, S.; Liu, J.; Zhu, X.; Madhyastha, H.; Chowdhury, M. Fedscale: Benchmarking model and system performance of federated learning at scale. In Proceedings of the International conference on machine learning, 2022; pp. 11814-11827.
196. Chai, D.; Wang, L.; Yang, L.; Zhang, J.; Chen, K.; Yang, Q. Fedeval: A holistic evaluation framework for federated learning. *arXiv preprint arXiv:2011.09655* **2020**.
197. Zhuang, W.; Wen, Y.; Zhang, X.; Gan, X.; Yin, D.; Zhou, D.; Zhang, S.; Yi, S. Performance optimization of federated person re-identification via benchmark analysis. In Proceedings of the Proceedings of the 28th ACM International Conference on Multimedia, 2020; pp. 955-963.
198. Chen, D.; Gao, D.; Kuang, W.; Li, Y.; Ding, B. pfl-bench: A comprehensive benchmark for personalized federated learning. *Advances in Neural Information Processing Systems* **2022**, *35*, 9344-9360.
199. He, C.; Balasubramanian, K.; Ceyani, E.; Yang, C.; Xie, H.; Sun, L.; He, L.; Yang, L.; Yu, P.S.; Rong, Y. Fedgraphnn: A federated learning system and benchmark for graph neural networks. *arXiv preprint arXiv:2104.07145* **2021**.
200. Caldas, S.; Duddu, S.M.K.; Wu, P.; Li, T.; Konečný, J.; McMahan, H.B.; Smith, V.; Talwalkar, A. Leaf: A benchmark for federated settings. *arXiv preprint arXiv:1812.01097* **2018**.
201. Luo, J.; Wu, X.; Luo, Y.; Huang, A.; Huang, Y.; Liu, Y.; Yang, Q. Real-world image datasets for federated learning. *arXiv preprint arXiv:1910.11089* **2019**.
202. Huang, X.; Han, L.; Li, D.; Xie, K.; Zhang, Y. A reliable and fair federated learning mechanism for mobile edge computing. *Computer Networks* **2023**, *226*, 109678.
203. Abdul Salam, M.; Fouad, K.M.; Elbably, D.L.; Elsayed, S.M. Federated learning model for credit card fraud detection with data balancing techniques. *Neural Computing and Applications* **2024**, *36*, 6231-6256.
204. Ducange, P.; Marcelloni, F.; Renda, A.; Ruffini, F. Federated Learning of XAI Models in Healthcare: A Case Study on Parkinson's Disease. *Cognitive Computation* **2024**, 1-26.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.