Computer Networks and Communications
https://ojs.wiserpub.com/index.php/cnc/

UNIVERSAL WISER PUBLISHER

Article

Survey: An Overview on Privacy Preserving Federated Learning in Health Data

Manzur Elahi[1,*], Hui Cui,[1,2] and Mohammed Kaosar[1]

[1] School of Information Technology, Murdoch University, Perth, Australia
[2] Faculty of Information Technology, Monash University, Melbourne, Australia
E-mail: elahi01@yahoo.com

Received: 2 October 2022; **Revised:** 6 December 2022; **Accepted:** 21 December 2022

Abstract: Machine learning now confronts two significant obstacles: the first is data isolation in most organizations' silos, and the second is data privacy and security enforcement. The widespread application of Machine Learning techniques in patient care is currently hampered by limited dataset availability for algorithm training and validation due to the lack of standardised electronic medical records and strict legal and ethical requirements to protect patient privacy. To avoid compromising patient privacy while supporting scientific analysis on massive datasets to improve patient care, it is necessary to analyse and implement Machine Learning solutions that fulfil data security and consumption demands. In this survey paper, we meticulously explain the existing works of federated learning from many perspectives to give a thorough overview and promote future research in this area. Then, we determine the current challenges, attack vectors and potential prospects for federated learning research. We analysed the similarities, differences and advantages between federated learning and other machine learning techniques. We also discussed about system and statistical heterogeneity and related efficient algorithms.

***Keywords*:** federated learning, homomorphic encryption, sensitive data, machine learning, artificial intelligence

1. Introduction

We currently live in a data-driven world, where most apps and services, such as healthcare and medical services, self-driving cars, and financial applications, are powered by artificial intelligence (AI) technology and complicated data-hungry Machine Learning (ML) algorithms. Furthermore, AI technology has yet to achieve its full potential, and long-standing issues have plagued the implementation of such AI/ML-based applications, one of which is centralised storage and processing [1].

In most real-world circumstances, data, particularly personal data, is created and kept, either on end-user devices or in data centres operated by service providers. Traditional ML algorithms are run in a centralised method, requiring training data to be combined in a data server. Large-scale data collecting and processing at a powerful cloud-based server in standard ML algorithms involves a single point of failure and the possibility of major data breaches. Centralised data processing and administration impose limited transparency and provenance on the system, which may result in a lack of confidence from end users as well as trouble complying with the regulatory authority such as EU General Data Protection Regulation (GDPR) [2], Australian Privacy Act (APA).

To overcome such challenges, Federated Learning (FL) was first presented by Google in 2016 [1]. FL is an algorithmic method that enables the training of ML models by delivering copies of a model to the location where data performs the training at the edge, minimizing the need to transport vast volumes of data to a central server for training. The data stays on the clients' originating devices, which receive a replica of the global model from the central server. Each device's data is used to train a local replica of the global model. Local training is

DOI: https://doi.org/10.37256/cnc.1120231992

used to update the model weights, and then the local copy is transmitted back to the central server. When the server gets the revised model, it aggregates the changes to improve the global model without disclosing any of the private data used to train it.

When compared to standard ML techniques, FL inherently provides advantages in terms of privacy. However, FL systems are sensitive to inference attacks (i.e., membership and reconstruction attacks) [3]. Malicious local participants can manipulate the training progression by delivering arbitrary updates to poison the global model [4,5].

Patient records are strictly confidential. As an example, pregnancy data is highly sensitive, particularly information linked to pregnancy termination, young women abortion. Often pregnant mother is unwilling to share pregnancy record/information to Spouse, Parents, family, and friends for various concerns. Privacy breach may result in catastrophic consequences such as domestic violence, shame, mental illness, family breakdown and social isolation etc.

Potential of Complex Data: Data sources in complex format such as Time-series measurements (fetal Heart rate or actigraphy), imaging data (ultrasound), and free text (e.g., diagnosis records, patient narratives). ML can provide information critical for a more complete understanding of maternal and neonatal health [6]. In WA Health Data Linkage Benefits [7], the Western Australian Data Linkage System explains how to design and implement a comprehensive, population-based system. Among the benefits are improvements in health policy and practice, community development, patient privacy preservation, economic stimulation of the research sector, database quality upgrade, and research cost reduction. In Pregnancy Data Driven Model [6] the Machine-learning methods were examined to provide deeper biological insights into a pregnancy that current methodologies have not revealed; to clarify the pathologies that affect a pregnancy; and to recommend the best approaches to address disparities in outcomes affecting vulnerable population. In this survey paper we will discuss about Federated Learning definitions, workflow, security risks and different privacy mechanisms, performance evaluation based on privacy protection mechanism, FL algorithms, statistical and system heterogeneity and efficient algorithm addressing those heterogeneity.

2. Related Works

Homogeneous patients with comparable traits are grouped to construct a high-performance model, and each group develops its own local and global models. The authors in [8] present an in-depth analysis of the previously published studies on FL as well as definitions, architectures, and applications. The survey [9] provides a detailed and organised analysis of Privacy Preserving Federated Learning (PPFL). They suggested a scenario-based taxonomy focused on the 5Ws (Who, What, When, Where, Why). They reviewed current approaches, examined privacy leakage issues in the FL from various angles, and identified potential research areas. Additionally, they summarised the interconnections between the proposed taxonomy and the practises now in use. The authors of the survey [1] perform an analysis of recent FL studies that focuses on GDPR compliance and privacy-preserving strategies. Authors in [10] suggests a federated transfer learning strategy for wearable medical equipment. Lyu et al. in [11] listed the dangers that threaten FL systems. The authors' major attention was directed against inference and poisoning attacks, which disrupt the expected model behaviour. However, the authors did not mention any of privacy-preserving mechanisms that could be used in the FL specially for Model and Data protection. Brief reviews of privacy-preserving FL mechanisms were provided by Li et al. [12] with an emphasis on homomorphic encryption, secure multiparty computing, and differential privacy. FL is used in [13] to address patient record privacy concerns. However, the authors assert that it will be challenging to accomplish a trade-off between what the model is learning globally considering local information from each data source due to the vast number of data sources with various volumes of data having multiple features. Because of this, the authors suggest a novel approach called FADL, in which the first layer of the neural network model is trained locally in each data source. The authors demonstrated the Neural Network model for ICU mortality prediction. In contrast, the following layers are trained in an FL manner utilising data from all sources. Their proposal surpasses the use of the standard FL for dispersed electronic health records and exhibits accuracy comparable to a centralised analysis. The future of digital health with FL [14] explores the federated future for digital health, providing context and information to the community on the benefits and consequences of FL for medical applications, as well as outlining important difficulties and barriers to implementing FL for digital health. Dayan, Roth, H.R., Zhong, A. et al. in [15] discussed about FL for COVID-19 Oxygen Prediction Patient Care “EXAM” AI Model. The prediction model was implemented in TensorFlow using the NVIDIA CLARA. The authors in [16] suggests a homomorphic encryption-based federated learning method for medical data.

Table 1 below shows that most research papers have covered some or either of FL characteristics. However, there needs to be more discussion about Health records Heterogeneity and algorithms. In this research paper, we discussed a wide variety of FL features, including FL definitions, security and privacy issues, different algorithms, System Heterogeneity, and Statistical Heterogeneity in Health Records. Moreover, we discussed the best performing algorithm for Heterogeneous Federated Learning. The main contributions of this survey are providing an overview and architecture of FL, discussion of FL based on data partitioning, protection mechanism, privacy risks. We highlighted FL algorithms comparison, practical implementation of FL based on medical records, performance impact and trade-off between security/privacy protection and performance/accuracy in FL (section 7). We also discussed efficient FL algorithms addressing both Statistical and System Heterogeneity in section 6.

Table 1. Overview and comparison of related work on FL

Ref	Summary of Contribution	Security and Privacy	FL Algorithm	Health	System and Statistical Heterogeneity
[1]	PPFL survey from the GDPR Perspective	Yes	Yes	No	No
[8]	FL definitions, architectures, and applications	Yes	No	No	No
[9]	Survey of PPFL, Review, Scenario-based taxonomy, and Future Directions	Yes	No	No	No
[10]	Federated transfer learning strategy for wearable medical equipment	No	Yes	Yes	No
[11]	FL threats and attacks	Yes	No	No	No
[12]	FL Privacy preserving mechanism	Yes	Yes	No	No
[13]	FL Algorithm for Distributed Electronic Health Record and mortality prediction	No	Yes	Yes	No
[14]	Benefits and challenges of FL for medical applications	Yes	No	Yes	No
[15]	CLARA FL for COVID-19 Oxygen prediction model	No	Yes	Yes	No
[16]	Homomorphic Encryption and FL based COVID-19 Detection	Yes	Yes	Yes	No

3. Preliminaries

3.1 *Federated Learning Overview*

Google was the first to propose a federated learning framework in 2016 [17,18]. The basic concept behind Google's approach is to construct machine-learning models based on datasets spread across several devices while limiting data sharing. Current enhancements focus on overcoming statistical problems [19,20] and improving privacy [21,22] in Federated Learning; there are several other initiatives to enhance federated learning personalization increase through research [23,19].

3.2 *Federated Learning Workflow Cycle*

Real-world FL applications such as Google's G-board have concentrated on Centralised FL frameworks, where the centralized control is performed by a Coordination orchestration server by requesting and aggregating training outcomes to/from local nodes. Decentralized FL is also an option, in which local nodes directly communicate their training results on a peer-to-peer basis [24]. The FL workflow cycle is divided into four steps (shown in Figure 1) [1].

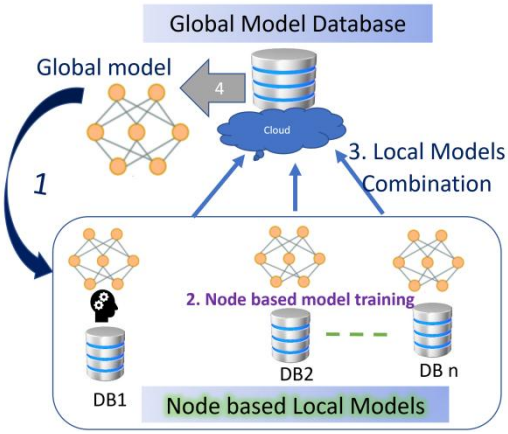

Figure 1. Federated Learning workflow cycle.

1) Participant node selection and global model transmission

The centralised global model server chooses a group of site/nodes who meet the conditions to participate in the training process and then broadcasts to the participants a global ML model (or global model changes) for the next training phase.

2) Node local model training

The participant nodes update their current local ML model after obtaining the global ML model from the server, and then train the new model using the device's local dataset. This stage is performed at local nodes, and it involves the installation of a FL client application on end-user devices to run training algorithms like Federated SGD and Federated Averaging, as well as to receive global model updates and transfer local ML model parameters from/to the server.

3) Local Models Combination

To update the global ML model, the centralised global model server collects enough locally trained ML models from participants (the next step). To prevent the server from accessing individual ML model parameters, this aggregation mechanism is necessary to combine several privacy preserving strategies such as safe aggregation, differential privacy, and sophisticated encryption methods.

4) Centralised Global Model Update

Based on the pooled model parameters collected in step 3 (local models' combination), the server updates the current global ML model. In the upcoming training phase, participants will receive an updated global model.

This four-step cycle is continued until the global model is adequately precise.

3.3 *Privacy of Federated Learning*

One of the most important aspects of Federated Learning is privacy; and privacy-preserving procedures are meant to maximise data value while guaranteeing that the original data is not shared with other people or organisations. Every stage of the training and deployment process is a target for a variety of new assaults and threats, making FL systems vulnerable. Which is disused in Federated Learning Privacy Risks section. Multiple approaches exist for attackers to take advantage of FL system weaknesses. Before delivering them to the central server, an attacker, for instance, may purposefully damage training data or local model updates on clients' devices. The central server and the clients' devices communicate model updates, which attackers may intercept and replace with malicious model updates. Numerous academics have offered defences against FL attacks and vulnerabilities to address these security risks. In this part, we discuss and evaluate federated learning privacy preservation techniques. The Cryptographic techniques commonly used in FL privacy mainly consist of Homomorphic Encryption, Secret Sharing, and Secure Multiparty Computation (SMC).

3.3.1 Homomorphic Encryption

Homomorphic encryption allows calculations on encrypted data without needing the secret key to decrypt the ciphertext [25]. The computation's results are encrypted and can only be decrypted by the computation's requester. Furthermore, homomorphic encryption assures that the decrypted result is identical to that computed on the unencrypted dataset. Homomorphic encryption protects sensitive medical data. The authors of [26] proposed an online secure multiparty computation (SMC) using homomorphic encryption for exchanging

medical information with hospitals. The authors Bocu et al. [27] described the homomorphic encryption scheme connected to healthcare information system employing heart rate data. The suggested approach effectively handled the criteria for the secure data processing for the 500 patients to overcome network and storage challenges. In [28], researchers created an IoT based platform featuring homomorphic encryption to prevent privacy leaks and spoofing threats in chronic illness monitoring. The findings indicate that homomorphic encryption provides cost-effective and simple privacy for sensitive health information. Homomorphic encryption systems may be classified into three types based on the supporting operations (addition, multiplication): Partially Homomorphic Encryption (PHE) features single type of operation (addition or multiplications) with an infinite number of times. Somewhat Homomorphic Encryption (SWHE) features both addition and multiplication operations a limited number of times. Fully Homomorphic Encryption (FHE) features unlimited number of operations for an unlimited number of times [29]. The significant advantage is that FHE ensures security by stronger encryption. Authors in [30] proposed a multi-party privacy preserving machine learning framework based on Partially Homomorphic Encryption and Federated Learning

Even though homomorphic encryption provides individuals with strict privacy guarantees because the original data in plaintext is never divulged, there is a practical limit to conducting computations over cipher-text due to the enormous processing overhead [31]. Homomorphic encryption will undoubtedly result in certain performance issues, such as the increased overhead of the encryption and decryption processes, which will have a significant impact on training efficiency. The network topology, encryption/decryption key length, and key replacement frequency, among other factors, all have an impact on the final performance.

3.3.2 Secure Multiparty Computation (SMC)

SMC, often called Multi Party Computation (MPC) or privacy-preserving computation, was initially proposed by Yao in 1982 [32] and has since been improved by several researchers. SMC is a cryptographic system that allows distributed parties to compute an objective function jointly while keeping their data private. All participating parties' data is kept private and all parties only allowed to acquire the outcome. The only information each party can obtain from the computation is the result and its own inputs. Classical secret sharing such as Shamir's secret sharing [33,34] and Verifiable Secret Sharing (VSS) schemes [35] are the basis for most of the SMC protocols [1].

There are several advantages of SMC, including that no trusted third parties are required, the compromise between data utility and data privacy is abolished, and a high level of accuracy is obtained. The downsides are high communication costs and computational overhead [9].

3.3.3 Perturbation Technique

The main principle of perturbation technique is to introduce noise to the original data, resulting in statistical information generated from the perturbed data that is statistically indistinguishable from the original data. Differential privacy, Additive perturbation, and Multiplicative perturbation are three often used perturbation techniques. The fundamental composition theorem leads to differential privacy for repeated applications of additive noise techniques. The following steps make up the actual design process for a differentially private additive-noise mechanism that implements a given functionality: approximating the functionality by sequentially composing bounded-sensitivity functions, selecting additive noise's parameter settings, and conducting a privacy analysis of the resulting mechanism [36]. The authors in Differential Privacy-enabled Federated Learning for Sensitive Health Data [37] analysed the effect of differing degrees of privacy on the performance of the FL model and collect insights on the usefulness of FL with and without differentiated privacy in healthcare applications.

3.4 *Types of Federated Learning*

This section will discuss how to classify Federated Learning depending on the data partitioning and distribution characteristics. The feature and sample spaces of the participating nodes data sets may differ. We categorise federated learning into three categories: horizontally federated learning, vertically federated learning, and federated transfer learning, based on data distribution among participants in the feature space and sample ID space [8].

3.4.1 Horizontal Federated Learning

This is the scenario when the decentralized datasets have the same feature space, but different sample ID space. This is a common occurrence in the same field. Student records from various colleges, for example, often have the same feature space but a separate ID space. Two regional banks may have quite diverse user groups

from their separate regions, with a relatively limited intersection set of their users. However, because their businesses are so similar, the feature spaces are the same. A horizontal federated learning system typically assumes honest participants and security against an honest-but-curious server [21,38]. Aggregation at the server is more straightforward since all clients may utilise a single model; the FedAvg method is popular for aggregation. Another simple to comprehend example would be a dataset that only included lung cancer patient records from a particular hospital.

3.4.2 Vertical Federated Learning

When the decentralized datasets have same sample ID space, but they differ in feature space. This is a common occurrence in a variety of areas. Each person, for example, may have many data records in financial systems, hospitals, ecommerce markets etc. In [9] the authors suggested a vertical federated-learning strategy to build a privacy-preserving logistic regression model. Authors investigated the impact of entity resolution on learning performance and used Taylor approximation to the loss and gradient functions, allowing homomorphic encryption to be used for privacy-preserving calculations.

3.4.3 Transfer Federated Learning

It refers to the situation in which the decentralised datasets differ in terms of both sample ID and feature space. A recommendation model that learns from varied app/software usage behaviours of different users in different countries is an example of Transfer FL. Because they utilise various programmes on their devices, they produce independent decentralised datasets with no overlap in sample and feature space. The federated transfer learning framework for wearable healthcare - FedHealth [10] has described FL featuring customised models using transfer learning.

Table 2 summarises the FL data partition type. Horizontal FL has same feature space but different sample ID; Vertical FL has similar sample ID but different feature space; Transfer or Hybrid FL has both different sample ID and feature space. Transfer Federated Learning is more suitable for Heterogeneous Electronic Health Record dispersed across the different hospitals or applications.

Table 2. Federated Learning Data Partitioning

FL Data Partitioning	Sample ID	Feature Space
Horizontal	Different	Same
Vertical	Same	Different
Transfer/Hybrid	Different	Different

3.5 *FL and Related Machine Learning Comparison*

There are a few related Machine Learning concepts like FL, including Distributed Machine Learning, Mobile Edge Computing, Split Learning, Privacy Preserving Machine Learning. We are comparing FL with the similar Machine Learning (ML) concepts highlighted from author of [9] in summarised content. The comparison is for main features, common features and different features of FL: Table 3 for comparison with distributed machine learning, Table 4 for Mobile edge Computing ML, Table 5 for Split Learning, Table 6 for Conventional Privacy Preserving ML. The important advantage of FL privacy preserving distributed mechanism suitable for large heterogeneous medical record.

Table 3. Comparison between Distribute Machine Learning and Federated Learning

Distributed ML	Federated Learning	
Main Features	Common Features	Different Features
• Data-parallel scheme • Model-parallel scheme • Multiple local nodes & an aggregator • Acceleration for large-scale datasets • Homogeneous data	• Large scale data sets • Distributed Computing	• Privacy-preserving • May be no central aggregators • No exchange of datasets • Heterogeneous/homogeneous data

Table 4. Comparison between Mobile Edge Computing and Federated Learning

Mobile Edge Computing ML	Federated Learning	
Main Features	Common Features	Different Features
• Three-layer architecture: end-users, edge servers & a cloud server • Low-latency (e.g., augmented reality) • IoT devices (e.g., smartphones, Sensors)	• Privacy-preserving • Large-scale datasets • Distributed computing	• May be no servers • No exchange of datasets

Table 5. Comparison between Split Learning and Federated Learning

Split Learning	Federated Learning	
Main Features	Common Features	Different Features
• Model splitting design • Collaborative learning • Clients & a central server • Communication efficiency	• Large-scale datasets • No Datasets exchange	• Entire model training • May be no central aggregators • Privacy-preserving

Table 6. Comparison between Conventional Privacy Preserving ML and Federated Learning

Conventional Privacy Preserving ML	Federated Learning	
Main Features	Common Features	Different Features
• Usually centralized computing • Privacy-preserving machine learning	• Privacy-preserving • Neural network	• Distributed computing • Multiple participants • Multiple private datasets • Large-scale datasets

In summary, the main advantages of Federated Learning over other ML techniques are: Privacy preserved architecture, No exchange of datasets, suitability for large scale heterogeneous data and heterogenous platforms. As highlighted in Table 6: Electronic Health Records centralised ML and data administration impose limited transparency on the system, which may result in a lack of confidence from end users, hospitals, and nodes. As well as implies trouble complying with the regulatory authority such as EU General Data Protection Regulation (GDPR), Australian Privacy Act (APA).

3.6 *Federated Learning Privacy Risks*

FL provides a framework for collaboratively training a global model with datasets stored in different clients. But FL may not always ensure sufficient privacy protection [39]. Beyond the restrictions that are basic and FL-specific, the security of FL systems is crucial for creating networks where users may cooperate, learn, and, most importantly, trust. Every stage of the training and deployment process is a target for a variety of new assaults and threats, making FL systems susceptible. Multiple approaches exist for attackers to take advantage of FL system weaknesses. Before delivering them to the central server, an attacker, for instance, may purposefully damage training data or local model updates on clients' devices. Someone may intercept the model updates sent back and forth between the customers' devices and the central server and substitute malicious model updates for them [17].

The potential privacy leakage issue can be divided into six important perspectives [9] shown in Figure 2.

Attackers: There are two types of attackers, internal (nodes and central global model server) and external (model consumers and eavesdroppers). An example of internal malicious attackers: The intermediate training model updates (e.g., weights and gradients) and the final model are accessible to the client nodes and the central server. As a result, certain individuals and an honest-but-curious server might be internal malevolent opponents attempting to obtain access to confidential data. Other type of malicious attackers can be external: Model consumers can access the whole model and through API interface the query results can be accessed [40]. By intercepting the connection between the participants and the server, eavesdroppers can steal intermediate training updates or the final model.

Attack Type: Active and passive attacks are the two types of privacy attacks. FL active attacks seek to actively affect the training process and collect sensitive information about training datasets [41]. Passive attacks are described as those that use information or learn from a system without altering it.

Attack Phase: Training phase and inference phase are two primary phases for privacy leakage. The training phase involved mostly of computing the local gradient on the client, aggregating the global model on the server, transferring intermediate updates between the clients and the aggregator, and delivering the final model to the clients [42]. The inference phase is primarily concerned with determining how to deliver the query service for customers [41].

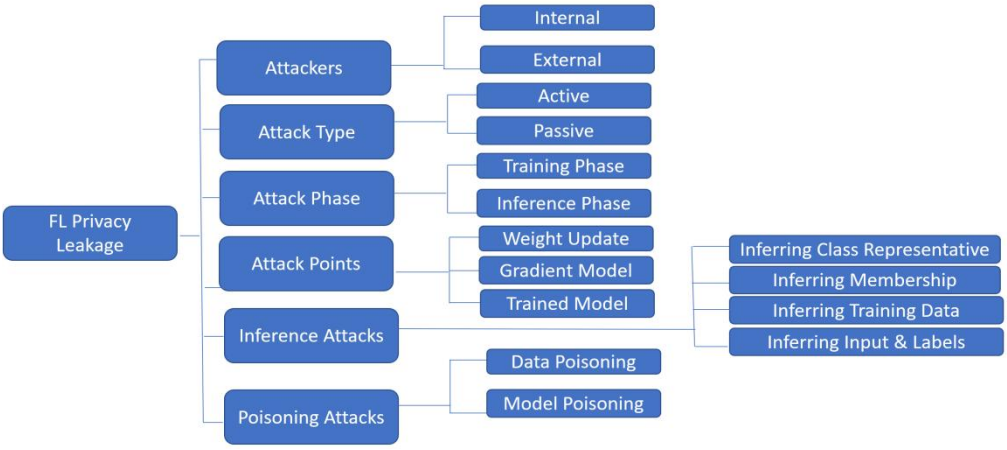

Figure 2. FL privacy leakage

Attack Points: There are three possible attack points; Weight update stage, gradient model stage and attack on trained model.

Inference Attacks: An inference privacy attack is typically used to gather sensitive information about training datasets. Inference attacks are classified into four types [43]: Inferring class representative, Inferring Membership, Inferring Training Data, Inferring Input and labels.

Poisoning Attacks: During poisoning attack, intruders manipulate the client's data or local model to influence the global model's performance/accuracy [44]. Mainly divided into Data and model poisoning. FL poisoning attack is a serious concern. Two types of FL poising attack shown in figure 3; consist of data poisoning (clean label and dirty label) and model poisoning (gradient manipulation, training rule manipulation, backdoor attacks).

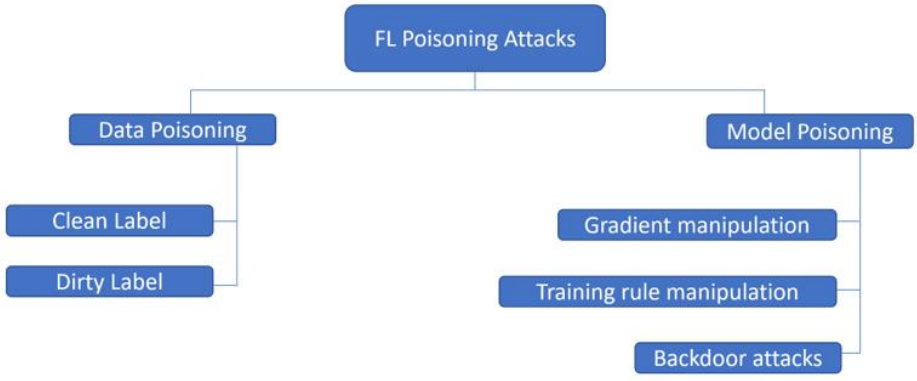

Figure 3. FL poisoning attack types

Figure 4 shows the poising attack effect or flow. Data poisoning may lead to model poisoning and result in a possible cause of poisoned model parameters. By sending poisoned local model updates to a coordination server, model poisoning attacks often try to manipulate the training process.

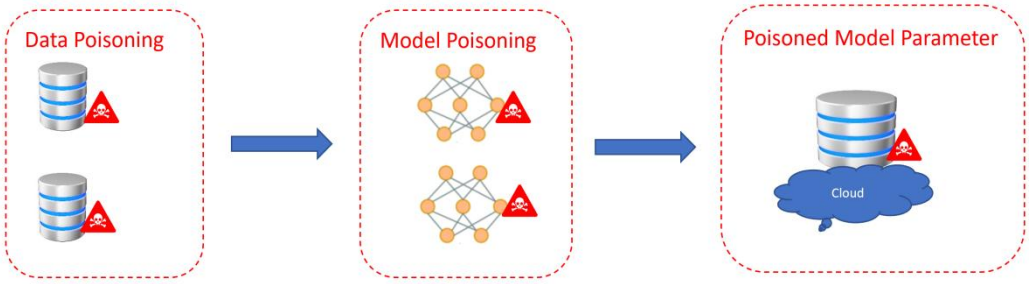

Figure 4. FL poisoning attack flow

4. Heterogeneity in Medical Records

The diversity of electronic health records (EHR) is inherent. Real-world FL deployments occur across a hospital network with various data samples, device capabilities, availability, network quality, security level, operating system platform, and database types. Such as in Department of Health Western Australia [45]; there are more than 400 clinical and enterprise business applications, with different platforms and databases, more than 540 sites across the state. According to the data structure, Electronic Health Records (EHR) can be categorised into three groups: structured data, unstructured data, and semi-structured data [46].

Structured data, such as numbers and symbols, can be represented by a consistent structure. Two-dimensional tables are used to express structured medical data that is kept in databases. A few examples include demographic data, an encounter, a diagnosis code, a procedure code, a laboratory result, a drug, an allergic reaction history, social data, etc.

Unstructured data includes both text and images. The fundamental purpose of an unstructured database is to store unstructured data. A clinical note, radiological report, surgical note, discharge note, etc., are some examples. It overcomes limitations of a fixed length of data and unchangeable definition of the relational database structure, permits repeated fields, subfields, and variable-length fields, and implements storage management.

Semi-structured data exists between structured and unstructured data. It is often self-descriptive; the structure and content of the data are blended with no evident distinctions. Reminiscent of resource description files, for example (RDF). Semi-structured data can be considered a subset of structured data, but its structure varies widely.

Heterogeneity thereby affects a federated network's performance and is demonstrable. This section breaks the heterogeneity into two primary categories—data diversity and node diversity.

Data Diversity/Non-IID Data: Many models for distributed optimization issues assume Independently and Identically Distributed data (IID). The difficulty of issue modelling, solution formulation, analysis, and optimization are increased because EHR contains many data points that differ significantly between silos. Statistical heterogeneity and non-IID make it challenging to obtain an ideal performance to learn a single, universally shared model.

Node diversity: Various health apps on a network may use different CPUs, RAM, network infrastructure, communication protocols, operating systems, cloud architectures, database diversity, etc. In conclusion, each node in the network has different processing, storage, and communication capabilities; additionally, FL networks must deal with stragglers.

5. ML Benefits and Barriers in Electronic Health Records

Since patient privacy is protected, it is incredibly challenging to gather medical data dispersed across many institutions. Medical data thus becomes a limited resource. The allocation of medical resources and illness diagnostics have undergone revolutionary changes due to the development of artificial intelligence.

Some of the benefits include: Live ML data/model synchronisation and sharing among other applications, facilities, agencies. ML can assist with Robust diagnosis, Interactive Alert generation, Realtime Data analysis without giving away the data, Disease prediction, Disease prevention, Forecasting and trend, Future planning, resource allocation, budgeting etc. The FL system is a desirable alternative, given the present emphasis on healthcare systems to improve sustainability. With FL, creating more universal models to help with clinical treatment is feasible. Less time spent in emergency rooms and hospitals, a slowing of the course of chronic

diseases, and better results sooner may be the effects of more effective and targeted patient care. FL has the potential to lower healthcare system expenses, making the healthcare system more sustainable [14]. Participants stand to gain from the FL system through the formation of cooperative partnerships and the benefits of information and expertise exchange. These frameworks represent the pinnacle of open innovation since they incorporate several outside stakeholders, including physicians, healthcare workers, administrators of healthcare facilities, data scientists, and software developers, who will contribute their skills. Manufacturers of healthcare gear and software might also gain from FL since it can help them continuously validate or enhance their ML-based systems by pooling the learning from several devices and apps without disclosing patient-specific information [14]. Introducing FL systems will result in lasting alliances between hospitals, medical facilities, and governments. FL Nodes or Participants stand to gain from the FL system through the formation of cooperative partnerships and the benefits of information and skill sharing. The ML frameworks represent the pinnacle of open innovation since they incorporate several outside stakeholders, including physicians, healthcare workers, administrators of healthcare facilities, data scientists, instrument manufacturers, and software developers, who will contribute and collaborate their skills. Introducing FL systems will result in lasting alliances between participant hospitals, medical facilities, and governments [47].

Some of ML/FL barriers: Isolated patient care applications, Data sharing restrictions and difficulties. There are strict legal and ethical requirements and standards to protect patient privacy such as Australian Privacy Act (APA), General Data Protection Regulation (GDPR). The standards are essential for data security and privacy reasons. However, they can make it challenging to communicate and link healthcare information between medical centres, hospitals, organisations, and governments. As a result, Electronic Health Records are frequently restricted to specific institutions, making it challenging to analyse and use deep learning in healthcare. There is a limited dataset availability for algorithm training and validation, Lack of standardised electronic medical records. When combined with other unidentified datasets, data biases from certain institutions might lead to models that scale poorly and underperform. Utilising data from many healthcare organisations while learning and improving combined models is one method for obtaining sufficiently vast and varied datasets.

When gathered by a centralised authority or company, personal data from sources like patient records, diagnostic devices, smartphones, wearables, and many others are sensitive and vulnerable to data breaches and misuse. Federated Learning is a decentralised machine learning approach that seeks to learn the same global machine learning model without using a central server. Instead, FL relies primarily on distributed calculations performed by the devices, with each user's data never leaving the node from which it originated. While significant progress has been achieved in this direction, important issues in terms of privacy, algorithmic efficiency, resilience, and scalability need to be solved.

6. FL Algorithms

We chose the top six open-source FL algorithms for centralised aggregation for this paper. Table 7 represents these six algorithms in brief comparison based on Non-IID support (Independently and Identically Distributed data) or Statistical Heterogeneity, Node Heterogeneity support, Equality (various users' data distribution may be considered equally without interfering with irrelevant elements) and communication efficiency. The authors of [48] discussed extremely popular and widely used FL algorithms, including their characteristics and limitation. We summarize couple of FL algorithms in Table 7.

Table 7. FL Algorithm comparison

FL Algorithms	Non IID/Statistical Heterogeneity	System/Node Heterogeneity	Impartiality	Communication Efficiency
FedAvg [49]	Yes	No	No	Yes [49]
FedProx [50]	Yes	Yes	Yes [50]	Yes
FedPD	Yes	No	No	No
SCAFFOLD	Yes	No	No	Yes
FeSEM	Yes	No	No	Yes

Federated Stochastic Expectation Maximization (*FeSEM*), a revolutionary multi-centre FL framework suggested by the authors in [51], updated several global models by combining data from various user groups.

Mainly, it is likely that the datasets created or derived from the same or an equal distribution for the users in the same group. The authors defined the multi-centre FL issue as the joint clustering of users followed by the global model optimisation for users in each cluster. The proposed multi-centre FL preserves the capacity to handle non-IID data on diverse datasets and inherits the federated SGD's efficient communication.

SCAFFOLD [52] proposes a Stochastic Controlled Averaging technique that leverages control variates (variance reduction) to offset the negative consequences of data heterogeneity. The algorithm calculates the update direction for each node and utilises the difference to fix the local update. The technique is said to overcome statistical heterogeneity issues and converge in far fewer communication cycles.

FedPD [53] introduced a novel algorithm design technique based on primal-dual optimisation. Federated Primal-Dual (FedPD) is a novel technique that seeks to deal with the general non-convex objective while obtaining the most significant possible optimisation and communication complexity for non-IID data.

FedAvg algorithm is the Federated Averaging method that combines model averaging on a server with local stochastic gradient descent (SGD) on each client. The authors in [49] conduct extensive tests on this technique, showing it can train deep networks on decentralised data with orders of magnitude fewer rounds of transmission and is resilient to imbalanced and non-IID data distributions. Since FedAvg trains high-quality models with only a small number of communication cycles, the studies demonstrate that federated learning may be made practical.

FedProx algorithm developed as the crucial realisation for the interaction between systems and statistical heterogeneity in federated learning. FedAvg does not let participating devices execute varying amounts of local work depending on their underlying system restrictions in the setting of systems heterogeneity; instead, it is typical to eliminate devices that are unable to compute. The problems presented by node and statistical heterogeneity are addressed by the FedProx algorithm conceptually and experimentally. The convergence behaviour of federated optimisation in heterogeneous networks can be negatively impacted by implicitly increasing statistical heterogeneity and eliminating stragglers (as in FedAvg) or naively absorbing partial information from stragglers. To address this problem, FedProx strengthens the method's stability by including a proximal term to the objective. FedProx enables different amounts of work to be completed locally among devices depending on their available system resources. It then aggregates the partial solutions supplied by the stragglers (as compared to dropping these devices) [50].

7. FL Performance Evaluation

Federated learning ML is proposed to safeguard sensitive patient information during multi-party computing operations. By avoiding the data from being centralised and susceptible, this strategy will increase the privacy and security of medical data. The authors in [16] proposed a Homomorphic encryption based model weight aggregation at the central server. With the help of the public patient record dataset for x-ray images, the training phase time performance of the suggested methods was evaluated.

The Figure 5 demonstrates the Federated Learning framework performance times in seconds for three alternative configurations: No model encryption, Model encryption utilising HE with ciphertext modulus=128, and Model encryption utilising HE with ciphertext modulus =192.

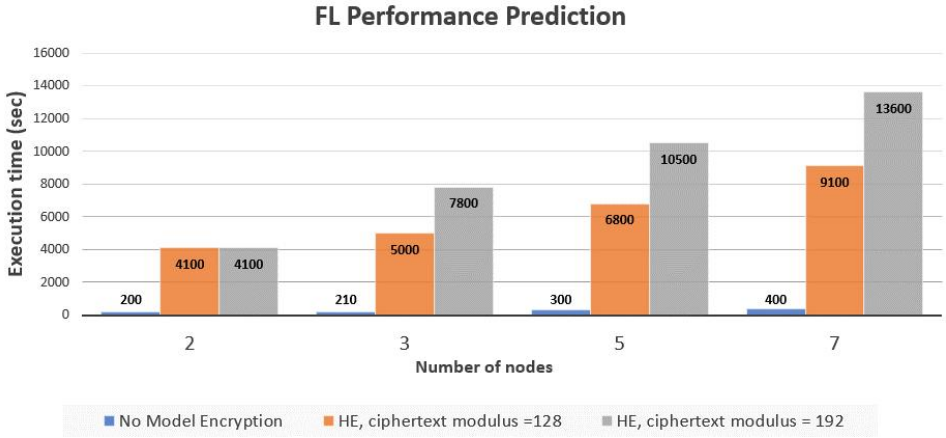

Figure 5. FL framework performance prediction

The execution time in the encrypted model is substantially greater than in the plain model. Due to the difficulty of homomorphic encryption and decrypting encrypted data, these exponential discrepancies exist. However, for two clients, the execution times of different ciphertext modulus values (12,192) are the same, but the difference in execution times increases as the number of clients increases. As a result, there will likely be a trade-off between model security and FL performance.

8. Conclusions

To provide a complete overview and encourage further study in this field, we painstakingly explain the current works of federated learning in this review article from various angles. We examined the parallels, contrasts, and benefits of federated learning compared to other machine learning approaches. We also identified the present issues, prospective threats, and future directions for federated learning research. The privacy attacks would cause huge destruction to the security and privacy of the medical data. This may limit the growth of improved patient and healthcare employing data-driven models. Therefore, it is important to take critical efforts to increase not just the safety of the information but also the manner data is handled.

Future Developments: In terms of future development, more powerful and scalable federated learning algorithms, such as vertical federated learning algorithms that partition features into various clients, should be addressed. Additionally, very efficient encryption methods (such as homomorphic encryption) will help to speed up learning. A robust privacy-protected learning algorithm, such as hybrid algorithms, anti-malicious attack clients' algorithm should be given greater consideration. In general, future research opportunities for sensitive medical records can focus on these two questions:

- How to ensure privacy and security for data and model for sensitive patient records in FL?
- How to achieve a more effective trade-off between security/privacy and performance/accuracy in FL?

Conflict of Interest

There is no conflict of interest for this study.

References

[1] Truong, N.; Sun, K.; Wang, S.; Guitton, F.; Guo, Y. Privacy preservation in federated learning: An insightf ul survey from the GDPR perspective. *Comput. Secur.* **2021**, *110*, 102402, https://doi.org/10.1016/j.cose.20 21.102402.

[2] Truong, N.B.; Sun, K.; Lee, G.M.; Guo, Y. GDPR-Compliant Personal Data Management: A Blockchain-Based Solution. *IEEE Trans. Inf. Forensic Security* **2020**, *15*, 1746–1761.

[3] Dwork, C., Smith, A., Steinke, T., Ullman, J., 2017. Exposed! A survey of attacks on private data. *Annu. Rev. Stat. Appl.* **2017**, *4*, 61–84, https://doi.org/10.1146/annurev-statistics-060116-054123.

[4] Bhagoji, A.N.; Chakraborty, S.; Mittal, P.; Calo, S.; Analyzing federated learning through an adversarial lens. In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 10–15 Junuary 2019. pp. 634–643.

[5] Fung, C.; Yoon, C.J.; Beschastnikh, I. Mitigating sybils in federated learning poisoning. *arXiv preprint* arXiv:1808.04866, https://doi.org/10.48550/arXiv.1808.04866.

[6] Espinosa, C.; Becker, M.; Marić, I.; Wong, R.J.; Shaw, G.M.; Gaudilliere, B.; Aghaeepour, N.; Stevenson, D.K. Data-Driven Modeling of Pregnancy-Related Complications. *Trends Mol. Med.* **2021**, *27*, 762–776, https://doi.org/10.1016/j.molmed.2021.01.007.

[7] Holman, C.D.J.; A Bass, J.; Rosman, D.L.; Smith, M.B.; Semmens, J.B.; Glasson, E.J.; Brook, E.L.; Trutwein, B.; Rouse, I.L.; Watson, C.R.; et al. A decade of data linkage in Western Australia: strategic design, applications and benefits of the WA data linkage system. Aust. Heal. Rev. **2008**, *32*, 766–777, https://doi.org/10.1071/ah080766.

[8] Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated Machine Learning: Concept and Applications. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–19, https://doi.org/10.1145/3298981.

[9] Yin, X.; Zhu, Y.; Hu, J. A Comprehensive Survey of Privacy-preserving Federated Learning. *ACM Comput. Surv.* **2021**, *54*, 1–36, https://doi.org/10.1145/3460427.

[10] Chen, Y.; Qin, X.; Wang, J.; Yu, C.; Gao, W. FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare. *IEEE Intell. Syst.* **2020**, *35*, 83–93, https://doi.org/10.1109/mis.2020.2988604.
[11] Lyu, L.; Yu, H.; Yang, Q. Threats to federated learning: A survey. *arXiv preprint* arXiv:2003.02133, https://doi.org/10.48550/arXiv.2003.02133.
[12] Li, Z.; Sharma, V.; Mohanty, S.P. Preserving Data Privacy via Federated Learning: Challenges and Solutions. *IEEE Consum. Electron. Mag.* **2020**, *9*, 8–16, https://doi.org/10.1109/mce.2019.2959108.
[13] Liu, D.; Miller, T.; Sayeed, R.; Mandl, K. FADL: Federated autonomous deep learning for distributed electronic health record. *arXiv preprint* arXiv:1811.11400, https://arxiv.org/abs/1811.11400.
[14] Rieke, N.; Hancox, J.; Li, W.; Milletarì, F.; Roth, H.R.; Albarqouni, S.; Bakas, S.; Galtier, M.N.; Landman, B.A.; Maier-Hein, K.; et al. The future of digital health with federated learning. *NPJ Digit. Med.* **2020**, *3*, 1–7, https://doi.org/10.1038/s41746-020-00323-1.
[15] Dayan, I.; Roth, H.R.; Zhong, A.; Harouni, A.; Gentili, A.; Abidin, A.Z.; Liu, A.; Costa, A.B.; Wood, B.J.; Tsai, C.-S.; et al. Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat. Med.* **2021**, *27*, 1735–1743, https://doi.org/10.1038/s41591-021-01506-3.
[16] Wibawa, F.; Catak, F.O.; Kuzlu, M.; Sarp, S.; Cali, U. Homomorphic Encryption and Federated Learning based Privacy-Preserving CNN Training: COVID-19 Detection Use-Case. In Proceedings of the 2022 European Interdisciplinary Cybersecurity Conference, Barcelona, Spain, 15–16 June, 2022. pp.85–90.
[17] Konecný, J.; McMahan, H.B.; Ramage, D.; Richtárik, P. Federated optimization: Distributed machine learning for on-device intelligence. *arXiv* arxiv:1610.02527, https://doi.org/10.48550/arXiv.1610.02527.
[18] McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B.A.y. Communication-efficient learning of deep networks from decentralized data. In Proceedings of 20th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL, USA, 9–11 May 2017. pp.1273–1282.
[19] Smith, V.; Chiang, C.-K.; Sanjabi, M.; Talwalkar, A.S. Federated multi-task learning. In Proceedings of the thirty-first Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017. pp. 4424–4434.
[20] Zhao, Y.; Li, M.; Lai, L.; Suda, N.; Civin, D.; Chandra, V. Federated Learning with Non-IID Data. *arXiv preprint* arXiv:1806.00582, https://doi.org/10.48550/arXiv.1806.00582.
[21] Bonawitz, K.; Ivanov, V.; Kreuter, B.; Marcedone, A.; McMahan, H.B.; Patel, S.; Ramage, D.; Segal, A.; Seth, K. Practical secure aggregation for privacy-preserving machine learning. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS'17). ACM, New York, NY, 1175–1191. https://doi.org/10.1145/3133956.3133982.
[22] Geyer, R.C.; Klein, T.; Nabi, M. Differentially private federated learning: A client level perspective. *arXiv preprint* arxiv:1712.07557, http://arxiv.org/abs/1712.
[23] Chen, F.; Dong, Z.; Li, Z.; He, X. Federated Meta-Learning with Fast Convergence and Efficient Communication. *arXiv preprint* arxiv:1802.07876, http://arxiv.org/abs/1802.07876.
[24] He, L.; Bian, A.; Jaggi, M. Cola: Decentralized linear learning. In Proceedings of the thirty-second Conference on Neural Information Processing Systems, Montreal, QC, Canada, 3–8 December, 2018. pp. 4536–4546.
[25] Gentry, C. Computing arbitrary functions of encrypted data. *Commun. ACM* **2010**, *53*, 97–105, https://doi.org/10.1145/1666420.1666444.
[26] Kumar, A.V.; Sujith, M.S.; Sai, K.T.; Rajesh, G.; Yashwanth, D.J.S. Secure Multiparty computation enabled E-Healthcare system with Homomorphic encryption. *IOP Conf. Series: Mater. Sci. Eng.* **2020**, *981*, 022079, https://doi.org/10.1088/1757-899x/981/2/022079.
[27] Bocu, R.; Costache, C. A homomorphic encryption-based system for securely managing personal health metrics data. *IBM J. Res. Dev.* **2018**, *62*, 1:1–1:10, https://doi.org/10.1147/jrd.2017.2755524.
[28] Talpur, M.S.H.; Alam Bhuiyan, Z.; Wang, G. Shared-node IoT network architecture with ubiquitous homo morphic encryption for healthcare monitoring. *Int. J. Embed. Syst.* **2015**, *7*, 43, https://doi.org/10.1504/ijes. 2015.066141.
[29] Acar, A., Aksu, H., Uluagac, A.S., Conti, M., 2018. A survey on homomorphic encryption schemes: Theory and implementation. *ACM Comput. Surv.* **2019**, *51*, 1–35, https://doi.org/10.1145/3214303.
[30] Fang, H.; Qian, Q. Privacy Preserving Machine Learning with Homomorphic Encryption and Federated Learning. *Futur. Internet* **2021**, *13*, 94, https://doi.org/10.3390/fi13040094.

[31] Gilad-Bachrach, R.; Dowlin, N.; Laine, K.; Lauter, K.; Naehrig, M.; Wernsing, J. Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy. In Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 19–24 June, 2016. pp. 201–210.
[32] Yao, A.C. Protocols for secure computations. In Proceedings of 23rd Annual Symposium on Foundations of Computer Science (sfcs 1982), Chicago, IL, USA, 3–5 November 1982. pp.160–164.
[33] Shamir, A. How to share a secret. *Commun. ACM* **1979**, *22*, 612–613, https://doi.org/10.1145/359168.359176.
[34] Brickell, E.F. Some ideal secret sharing schemes. In *Workshop on the Theory and Application of Cryptographic Techniques*, Springer: Berlin, Germany, 1989; pp. 468–475.
[35] Chor, B.; Goldwasser, S.; Micali, S.; Awerbuch, B. Verifiable secret sharing and achieving simultaneity in the presence of faults. In Proceedings of 26th Annual Symposium on Foundations of Computer Science (sfcs 1985), Portland, OR, USA, 21–23 October 1985. pp. 383–395.
[36] Abadi, M.; Chu, A.; Goodfellow, I.; McMahan, H.B.; Mironov, I.; Talwar, K.; Zhang, L. Deep Learning with Differential Privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016; pp. 308–318, https://doi.org/10.1145/2976749.2978318.
[37] Choudhury, O.; Gkoulalas-Divanis, A.; Salonidis, T.; Sylla, I.; Park, Y.; Hsu, G.; Das, A. Differential privacy-enabled federated learning for sensitive health data. *arXiv preprint* arXiv:1910.02578, https://doi.org/10.48550/arXiv.1910.02578.
[38] Phong, L.T.; Aono, Y.; Hayashi, T.; Wang, L.; Moriai, S. Privacy-Preserving Deep Learning via Additively Homomorphic Encryption. *IEEE Trans. Inf. Forensics Secur.* **2017**, *13*, 1333–1345, https://doi.org/10.1109/tifs.2017.2787987.
[39] Geiping, J.; Bauermeister, H.; Dröge, H.; Moeller, M. Inverting gradients-how easy is it to break privacy in federated learning. In Proceedings of the thirty-fourth Conference on Neural Information Processing Systems, Virtual Conference, 6–12 December 2020. pp.16937–16947.
[40] Truex, S.; Liu, L.; Gursoy, M.E.; Yu, L.; Wei, W. Demystifying Membership Inference Attacks in Machine Learning as a Service. *IEEE Trans. Serv. Comput.* **2019**, *14*, 2073–2089, https://doi.org/10.1109/tsc.2019.2897554.
[41] Nasr, M.; Shokri, R.; Houmansadr, A. Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In Proceedings of 2019 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 19–23 May 2019. pp.739–753.
[42] Hao, M.; Li, H.; Luo, X.; Xu, G.; Yang, H.; Liu, S. Efficient and Privacy-Enhanced Federated Learning for Industrial Artificial Intelligence. *IEEE Trans. Ind. Informatics* **2019**, *16*, 6532–6542, https://doi.org/10.1109/tii.2019.2945367.
[43] Enthoven D.; Al-Ars, Z. An overview of federated deep learning privacy attacks and eefensive strategies. *arXiv* arXiv:2004.04676, https://arxiv.org/abs/2004.04676.
[44] Benmalek, M., Benrekia, M.A., Challal, Y. Security of federated learning: Attacks, defensive mechanisms, and challenges. *Revue d'Intelligence Artificielle*, **2022**, *36*, 49–59, https://doi.org/10.18280/ria.360106.
[45] https://www.hss.health.wa.gov.au/Our-Services/Information-Communication-Technology (Accessed on 28 December 2022)
[46] Yue, L., Tian, D., Chen, W. et al. Deep learning for heterogeneous medical data analysis. *World Wide Web* **2020**, *23*, 2715–2737. https://doi.org/10.1007/s11280-019-00764-z.
[47] Long, G.; Shen, T.; Tan, Y.; Gerrard, L.; Clarke, A.; Jiang, J. Federated Learning for Privacy-Preserving Open Innovation Future on Digital Health. In *Humanity Driven AI*, Chen, F., Zhou, J., eds.; Springer: New York, NY, USA, 2022; pp.113–133, https://doi.org/10.1007/978-3-030-72188-6_6.
[48] Liu, J.; Huang, J.; Zhou, Y.; Li, X.; Ji, S.; Xiong, H.; Dou, D. From distributed machine learning to federated learning: a survey. Knowl. *Inf. Syst.* **2022**, *64*, 885–917, https://doi.org/10.1007/s10115-022-01664-x.
[49] McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B.A.y. Communication-efficient learning of deep networks from decentralized data. In Proceedings of 20th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL, USA, 9–11 May 2017. pp.1273–1282.

[50] Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated Optimization in Heterogeneous Networks. In Proceedings of the third Conference on Machine Learning and Systems, Austin, TX, USA, 2–4 March, 2020. pp.429–450.

[51] Long, G.; Xie, M.; Shen, T.; Zhou, T.; Wang, X.; Jiang, J. Multi-center federated learning: clients clustering for better personalization. *World Wide Web* **2022**, 1–20, https://doi.org/10.1007/s11280-022-01046-x.

[52] Karimireddy, S.P.; Kale, S.; Mohri, M.; Reddi, S.J.; Stich, S.U.; Suresh, A.T. SCAFFOLD: Stochastic Cont rolled Averaging for Federated Learning. *arXiv preprint* arXiv:1910.06378, https://arxiv.org/abs/1910.0637 8.

[53] Zhang, X.; Hong, M.; Dhople, S.; Yin, W.; Liu, Y. FedPD: A Federated Learning Framework with Optimal Rates and Adaptivity to Non-IID Data. *arXiv preprint* arXiv:2005.11418, https://arxiv.org/abs/2005.11418.