Journal of Healthcare Informatics Research (2021) 5:1–19
https://doi.org/10.1007/s41666-020-00082-4

RESEARCH ARTICLE

Federated Learning for Healthcare Informatics

Jie Xu[1] · Benjamin S. Glicksberg[2] · Chang Su[1] · Peter Walker[3] · Jiang Bian[4] · Fei Wang[1]

Received: 19 August 2020 / Revised: 21 October 2020 / Accepted: 30 October 2020 /
Published online: 12 November 2020

Abstract

With the rapid development of computer software and hardware technologies, more and more healthcare data are becoming readily available from clinical institutions, patients, insurance companies, and pharmaceutical industries, among others. This access provides an unprecedented opportunity for data science technologies to derive data-driven insights and improve the quality of care delivery. Healthcare data, however, are usually fragmented and private making it difficult to generate robust results across populations. For example, different hospitals own the electronic health records (EHR) of different patient populations and these records are difficult to share across hospitals because of their sensitive nature. This creates a big barrier for developing effective analytical approaches that are generalizable, which need diverse, "big data." Federated learning, a mechanism of training a shared global model with a central server while keeping all the sensitive data in local institutions where the data belong, provides great promise to connect the fragmented healthcare data sources with privacy-preservation. The goal of this survey is to provide a review for federated learning technologies, particularly within the biomedical space. In particular, we summarize the general solutions to the statistical challenges, system challenges, and privacy issues in federated learning, and point out the implications and potentials in healthcare.

Keywords Federated learning · Healthcare · Privacy

✉ Fei Wang
few2001@med.cornell.edu

[1] Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, USA

[2] Institute for Digital Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[3] U.S. Department of Defense Joint Artificial Intelligence Center, Washington, D.C., USA

[4] Department of Health Outcomes and Biomedical Informatics, College of Medicine, University of Florida, Gainesville, FL, USA

1 Introduction

The recent years have witnessed a surge of interest related to healthcare data analytics, due to the fact that more and more such data are becoming readily available from various sources including clinical institutions, patient individuals, insurance companies, and pharmaceutical industries, among others. This provides an unprecedented opportunity for the development of computational techniques to dig data-driven insights for improving the quality of care delivery [72, 105].

Healthcare data are typically fragmented because of the complicated nature of the healthcare system and processes. For example, different hospitals may be able to access the clinical records of their own patient populations only. These records are highly sensitive with protected health information (PHI) of individuals. Rigorous regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) [32], have been developed to regulate the process of accessing and analyzing such data. This creates a big challenge for modern data mining and machine learning (ML) technologies, such as deep learning [61], which typically requires a large amount of training data.

Federated learning is a paradigm with a recent surge in popularity as it holds great promise on learning with fragmented sensitive data. Instead of aggregating data from different places all together, or relying on the traditional discovery then replication design, it enables training a shared global model with a central server while keeping the data in local institutions where the they originate.

The term "federated learning" is not new. In 1976, Patrick Hill, a philosophy professor, first developed the Federated Learning Community (FLC) to bring people together to jointly learn, which helped students overcome the anonymity and isolation in large research universities [42]. Subsequently, there were several efforts aiming at building federations of learning content and content repositories [6, 74, 83]. In 2005, Rehak et al. [83] developed a reference model describing how to establish an interoperable repository infrastructure by creating federations of repositories, where the metadata are collected from the contributing repositories into a central registry provided with a single point of discovery and access. The ultimate goal of this model is to enable learning from diverse content repositories. These practices in federated learning community or federated search service have provided effective references for the development of federated learning algorithms.

Federated learning holds great promises on healthcare data analytics. For both provider (e.g., building a model for predicting the hospital readmission risk with patient Electronic Health Records (EHR) [71]) and consumer (patient)-based applications (e.g., screening atrial fibrillation with electrocardiograms captured by smartwatch [79]), the sensitive patient data can stay either in local institutions or with individual consumers without going out during the federated model learning process, which effectively protects the patient privacy. The goal of this paper is to review the setup of federated learning, discuss the general solutions and challenges, and envision its applications in healthcare.

In this review, after a formal overview of federated learning, we summarize the main challenges and recent progress in this field. Then we illustrate the potential of federated learning methods in healthcare by describing the successful recent research.

At last, we discuss the main opportunities and open questions for future applications in healthcare.

Difference with Existing Reviews There has been a few review articles on federated learning recently. For example, Yang et al. [109] wrote the early federated learning survey summarizing the general privacy-preserving techniques that can be applied to federated learning. Some researchers surveyed sub-problems of federated learning, e.g., personalization techniques [59], semi-supervised learning algorithms [49], threat models [68], and mobile edge networks [66]. Kairouz et al. [51] discussed recent advances and presented an extensive collection of open problems and challenges. Li et al. [63] conducted the review on federated learning from a system viewpoint. Different from those reviews, this paper provided the potential of federated learning to be applied in healthcare. We summarized the general solution to the challenges in federated learning scenario and surveyed a set of representative federated learning methods for healthcare. In the last part of this review, we outlined some directions or open questions in federated learning for healthcare. An early version of this paper is available on arXiv [107].

2 Federated Learning

Federated learning is a problem of training a high-quality shared global model with a central server from decentralized data scattered among large number of different clients (Fig. 1). Mathematically, assume there are K activated clients where the data reside in (a client could be a mobile phone, a wearable device, or a clinical institution data warehouse, etc.). Let $\mathcal{D}_k$ denote the data distribution associated with client k and

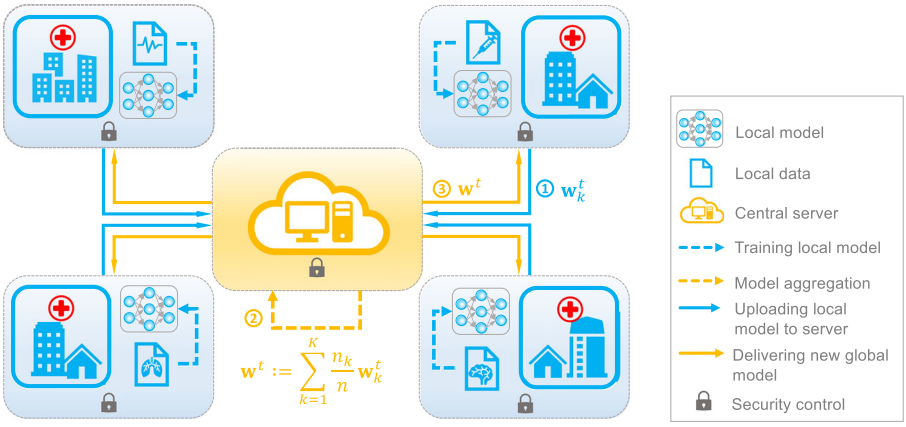

Fig. 1 Schematic of the federated learning framework. The model is trained in a distributed manner: the institutions periodically communicate the local updates with a central server to learn a global model; the central server aggregates the updates and sends back the parameters of the updated global model

n_k the number of samples available from that client. $n = \sum_{k=1}^{K} n_k$ is the total sample size. Federated learning problem boils down to solving a empirical risk minimization problem of the form [56, 57, 69]:

$$\min_{\mathbf{w}\in\mathbb{R}^d} F(\mathbf{w}) := \sum_{k=1}^{K} \frac{n_k}{n} F_k(\mathbf{w}) \quad \text{where} \quad F_k(\mathbf{w}) := \frac{1}{n_k} \sum_{\mathbf{x}_i \in \mathcal{D}_k} f_i(\mathbf{w}), \tag{1}$$

where $\mathbf{w}$ is the model parameter to be learned. The function f_i is specified via a loss function dependent on a pair of input-output data pair $\{\mathbf{x}_i, y_i\}$. Typically, $\mathbf{x}_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$ or $y_i \in \{-1, 1\}$. Simple examples include:

- linear regression: $f_i(\mathbf{w}) = \frac{1}{2}(\mathbf{x}_i^\top \mathbf{w} - y_i)^2$, $y_i \in \mathbb{R}$;
- logistic regression: $f_i(\mathbf{w}) = -\log(1 + \exp(-y_i \mathbf{x}_i^\top \mathbf{w}))$, $y_i \in \{-1, 1\}$;
- support vector machines: $f_i(\mathbf{w}) = \max\{0, 1 - y_i \mathbf{x}_i^\top \mathbf{w}\}$, $y_i \in \{-1, 1\}$.

In particular, algorithms for federated learning face with a number of challenges [13, 96], specifically:

- **Statistical Challenge:** The data distribution among all clients differ greatly, i.e., $\forall k \neq \tilde{k}$, we have $\mathbb{E}_{\mathbf{x}_i \sim \mathcal{D}_k}[f_i(\mathbf{w}; \mathbf{x}_i)] \neq \mathbb{E}_{\mathbf{x}_i \sim \mathcal{D}_{\tilde{k}}}[f_i(\mathbf{w}; \mathbf{x}_i)]$. It is such that any data points available locally are far from being a representative sample of the overall distribution, i.e., $\mathbb{E}_{\mathbf{x}_i \sim \mathcal{D}_k}[f_i(\mathbf{w}; \mathbf{x}_i)] \neq F(\mathbf{w})$.
- **Communication Efficiency:** The number of clients K is large and can be much bigger than the average number of training sample stored in the activated clients, i.e., $K \gg (n/K)$.
- **Privacy and Security:** Additional privacy protections are needed for unreliable participating clients. It is impossible to ensure all clients are equally reliable.

Next, we will survey, in detail, the existing federated learning related works on handling such challenges.

2.1 Statistical Challenges of Federated Learning

The naive way to solve the federated learning problem is through Federated Averaging (*FedAvg*) [69]. It is demonstrated can work with certain non independent identical distribution (non-IID) data by requiring all the clients to share the same model. However, *FedAvg* does not address the statistical challenge of strongly skewed data distributions. The performance of convolutional neural networks trained with *FedAvg* algorithm can reduce significantly due to the weight divergence [111]. Existing research on dealing with the statistical challenge of federated learning can be grouped into two fields, i.e., consensus solution and pluralistic solution.

2.1.1 Consensus Solution

Most centralized models are trained on the aggregated training samples obtained from the samples drawn from the local clients [96, 111]. Intrinsically, the centralized model is trained to minimize the loss with respect to the uniform distribution [73]: $\bar{\mathcal{D}} =$

$\sum_{k=1}^{K} \frac{n_k}{n} \mathcal{D}_k$, where $\bar{\mathcal{D}}$ is the target data distribution for the learning model. However, this specific uniform distribution is not an adequate solution in most scenarios.

To address this issue, the recent proposed solution is to model the target distribution or force the data adapt to the uniform distribution [73, 111]. Specifically, Mohri et al. [73] proposed a minimax optimization scheme, i.e., agnostic federated learning (AFL), where the centralized model is optimized for any possible target distribution formed by a mixture of the client distributions. This method has only been applied at small scales. Compared to AFL, Li et al. [64] proposed q-Fair Federated Learning (q-*FFL*), assigning higher weight to devices with poor performance, so that the distribution of accuracy in the network reduces in variance. They empirically demonstrate the improved flexibility and scalability of q-*FFL* compared to AFL.

Another commonly used method is globally sharing a small portion of data between all the clients [75, 111]. The shared subset is required containing a uniform distribution over classes from the central server to the clients. In addition to handle non-IID issue, sharing information of a small portion of trusted instances and noise patterns can guide the local agents to select compact training subset, while the clients learn to add changes to selected data samples, in order to improve the test performance of the global model [38].

2.1.2 Pluralistic Solution

Generally, it is difficult to find a consensus solution $\mathbf{w}$ that is good for all components $\mathcal{D}_i$. Instead of wastefully insisting on a consensus solution, many researchers choose to embracing this heterogeneity.

Multi-task learning (MTL) is a natural way to deal with the data drawn from different distributions. It directly captures relationships among non-IID and unbalanced data by leveraging the relatedness between them in comparison to learn a single global model. In order to do this, it is necessary to target a particular way in which tasks are related, e.g., sharing sparsity, sharing low-rank structure, and graph-based relatedness. Recently, Smith et al. [96] empirically demonstrated this point on real-world federated datasets and proposed a novel method *MOCHA* to solve a general convex MTL problem with handling the system challenges at the same time. Later, Corinzia et al. [22] introduced *VIRTUAL*, an algorithm for federated multi-task learning with non-convex models. They consider the federation of central server and clients as a Bayesian network and perform training using approximated variational inference. This work bridges the frameworks of federated and transfer/continuous learning.

The success of multi-task learning rests on whether the chosen relatedness assumptions hold. Compared to this, pluralism can be a critical tool for dealing with heterogeneous data without any additional or even low-order terms that depend on the relatedness as in MTL [28]. Eichner et al. [28] considered training in the presence of block-cyclic data and showed that a remarkably simple pluralistic approach can entirely resolve the source of data heterogeneity. When the component distributions are actually different, pluralism can outperform the "ideal" IID baseline.

2.2 Communication Efficiency of Federated Learning

In federated learning setting, training data remain distributed over a large number of clients each with unreliable and relatively slow network connections. Naively for synchronous protocol in federated learning [58, 96], the total number of bits that required during uplink (clinets → server) and downlink (server → clients) communication by each of the K clients during training is given by:

$$\mathcal{B}^{\text{up/down}} \in \mathcal{O}(U \times \underbrace{|\mathbf{w}| \times (H(\triangle \mathbf{w}^{\text{up/down}}) + \beta)}_{\text{update size}}) \tag{2}$$

where U is the total number of updates performed by each client, $|\mathbf{w}|$ is the size of the model and $H(\triangle \mathbf{w}^{\text{up/down}})$ is the entropy of the weight updates exchanged during transmitting process. β is the difference between the true update size and the minimal update size (which is given by the entropy) [89]. Apparently, we can consider three ways to reduce the communication cost: (a) reduce the number of clients K, (b) reduce the update size, (c) reduce the number of updates U. Starting at these three points, we can organize existing research on communication-efficient federated learning into four groups, i.e., model compression, client selection, updates reducing, and peer-to-peer learning (Fig. 2).

2.2.1 Client Selection

The most natural and rough way for reducing communication cost is to restrict the participated clients or choose a fraction of parameters to be updated at each round. Shokri et al. [92] use the selective stochastic gradient descent protocol, where the

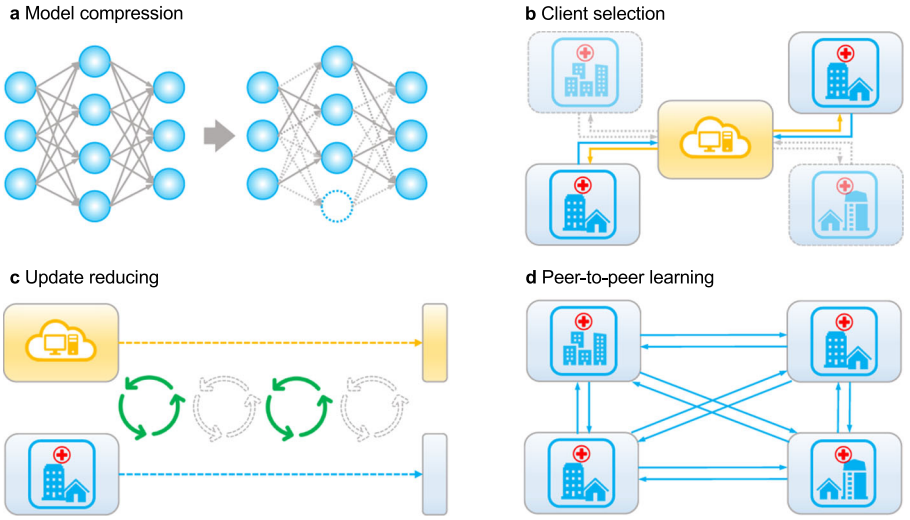

Fig. 2 Communication efficient federated learning methods. Existing research on improving communication efficiency can be categorized into **a** model compression, **b** client selection, **c** updates reducing, and **d** peer-to-peer learning

selection can be completely random or only the parameters whose current values are farther away from their local optima are selected, i.e., those that have a larger gradient. Nishio et al. [75] proposed a new protocol referred to as *FedCS*, where the central server manages the resources of heterogeneous clients and determines which clients should participate the current training task by analyzing the resource information of each client, such as wireless channel states, computational capacities, and the size of data resources relevant to the current task. Here, the server should decide how much data, energy, and CPU resources used by the mobile devices such that the energy consumption, training latency, and bandwidth cost are minimized while meeting requirements of the training tasks. Anh [5] thus proposes to use the Deep Q-Learning [102] technique that enables the server to find the optimal data and energy management for the mobile devices participating in the mobile crowd-machine learning through federated learning without any prior knowledge of network dynamics.

2.2.2 Model Compression

The goal of model compression is to compress the server-to-client exchanges to reduce uplink/downlink communication cost. The first way is through structured updates, where the update is directly learned from a restricted space parameterized using a smaller number of variables, e.g., sparse, low-rank [58], or more specifically, pruning the least useful connections in a network [37, 113], weight quantization [17, 89], and model distillation [43]. The second way is lossy compression, where a full model update is first learned and then compressed using a combination of quantization, random rotations, and subsampling before sending it to the server [2, 58]. Then the server decodes the updates before doing the aggregation.

Federated dropout, in which each client, instead of locally training an update to the whole global model, trains an update to a smaller sub-model [12]. These sub-models are subsets of the global model and, as such, the computed local updates have a natural interpretation as updates to the larger global model. Federated dropout not only reduces the downlink communication but also reduces the size of uplink updates. Moreover, the local computational costs is correspondingly reduced since the local training procedure dealing with parameters with smaller dimensions.

2.2.3 Updates Reduction

Kamp et al. [52] proposed to average models dynamically depending on the utility of the communication, which leads to a reduction of communication by an order of magnitude compared to periodically communicating state-of-the-art approaches. This facet is well suited for massively distributed systems with limited communication infrastructure. Bui et al. [11] improved federated learning for Bayesian neural networks using partitioned variational inference, where the client can decide to upload the parameters back to the central server after multiple passes through its data, after one local epoch, or after just one mini-batch. Guha et al. [35] focused on techniques for one-shot federated learning, in which they learn a global model from data in the network using only a single round of communication between the devices and

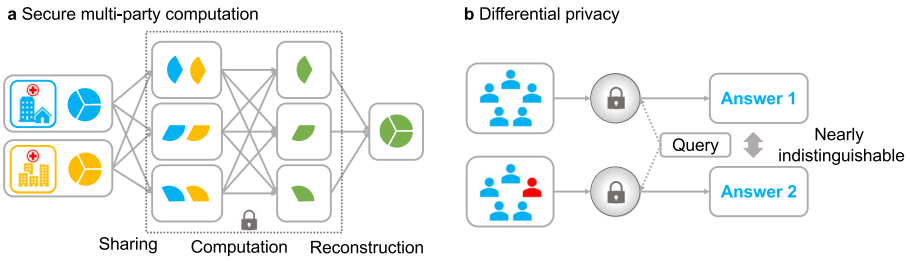

Fig. 3 Privacy-preserving schemes. **a** Secure multi-party computation. In security sharing, security values (blue and yellow pie) are split into any number of shares that are distributed among the computing nodes. During the computation, no computation node is able to recover the original value nor learn anything about the output (green pie). Any nodes can combine their shares to reconstruct the original value. **b** Differential privacy. It guarantees that anyone seeing the result of a differentially private analysis will make the same inference (answer 1 and answer 2 are nearly indistinguishable)

the central server. Besides above works, Ren et al. [84] theoretically analyzed the detailed expression of the learning efficiency in the CPU scenario and formulate a training acceleration problem under both communication and learning resource budget. Reinforcement learning and round robin learning are widely used to manage the communication and computation resources [5, 46, 106, 114].

2.2.4 Peer-to-Peer Learning

In federated learning, a central server is required to coordinate the training process of the global model. However, the communication cost to the central server may be not affordable since a large number of clients are usually involved. Also, many practical peer-to-peer networks are usually dynamic, and it is not possible to regularly access a fixed central server. Moreover, because of the dependence on central server, all clients are required to agree on one trusted central body, and whose failure would interrupt the training process for all clients. Therefore, some researches began to study fully decentralized framework where the central server is not required [41, 60, 85, 91]. The local clients are distributed over the graph/network where they only communicate with their one-hop neighbors. Each client updates its local belief based on own data and then aggregates information from the one-hop neighbors.

2.3 Privacy and Security

In federated learning, we usually assume the number of participated clients (e.g., phones, cars, clinical institutions...) is large, potentially in the thousands or millions. It is impossible to ensure none of the clients is malicious. The setting of federated learning, where the model is trained locally without revealing the input data or the model's output to any clients, prevents direct leakage while training or using the model. However, the clients may infer some information about another client's private dataset given the execution of $f(\mathbf{w})$, or over the shared predictive model $\mathbf{w}$ [100]. To this end, there have been many efforts focus on privacy either from an individual

point of view or multiparty views, especially in social media field which significantly exacerbated multiparty privacy (MP) conflicts [97, 98] (Fig. 3).

2.3.1 Secure Multi-party Computation

Secure multi-party computation (SMC) has a natural application to federated learning scenarios, where each individual client uses a combination of cryptographic techniques and oblivious transfer to jointly compute a function of their private data [8, 78]. Homomorphic encryption is a public key system, where any party can encrypt its data with a known public key and perform calculations with data encrypted by others with the same public key [29]. Due to its success in cloud computing, it comes naturally into this realm, and it has certainly been used in many federated learning researches [14, 40].

Although SMC guarantees that none of the parties shares anything with each other or with any third party, it can not prevent an adversary from learning some individual information, e.g., which clients' absence might change the decision boundary of a classifier, etc. Moreover, SMC protocols are usually computationally expensive even for the simplest problems, requiring iterated encryption/decryption and repeated communication between participants about some of the encrypted results [78].

2.3.2 Differential Privacy

Differential privacy (DP) [26] is an alternative theoretical model for protecting the privacy of individual data, which has been widely applied to many areas, not only traditional algorithms, e.g., boosting [27], principal component analysis [15], support vector machine [86], but also deep learning research [1, 70]. It ensures that the addition or removal does not substantially affect the outcome of any analysis and is thus also widely studied in federated learning research to prevent the indirect leakage [1, 70, 92]. However, DP only protects users from data leakage to a certain extent and may reduce performance in prediction accuracy because it is a lossy method [18]. Thus, some researchers combine DP with SMC to reduce the growth of noise injection as the number of parties increases without sacrificing privacy while preserving provable privacy guarantees, protecting against extraction attacks and collusion threats [18, 100].

3 Applications

Federated learning has been incorporated and utilized in many domains. This widespread adoption is due in part by the fact that it enables a collaborative modeling mechanism that allows for efficient ML all while ensuring data privacy and legal compliance between multiple parties or multiple computing nodes. Some promising examples that highlight these capabilities are virtual keyboard prediction [39, 70], smart retail [112], finance [109], and vehicle-to-vehicle communication [88]. In this section, we focus primarily on applications within the healthcare space and also

discuss promising applications in other domains since some principles can be applied to healthcare.

3.1 Healthcare

EHRs have emerged as a crucial source of real world healthcare data that has been used for an amalgamation of important biomedical research [30, 47], including for machine learning research [72]. While providing a huge amount of patient data for analysis, EHRs contain systemic and random biases overall and specific to hospitals that limit the generalizability of results. For example, Obermeyer et al. [76] found that a commonly used algorithm to determine enrollment in specific health programs was biased against African Americans, assigning the same level of risk to healthier Caucasian patients. These improperly calibrated algorithms can arise due to a variety of reasons, such as differences in underlying access to care or low representation in training data. It is clear that one way to alleviate the risk for such biased algorithms is the ability to learn from EHR data that is more representative of the global population and which goes beyond a single hospital or site. Unfortunately, due to a myriad of reasons such as discrepant data schemes and privacy concerns, it is unlikely that data will eve be connected together in a single database to learn from all at once. The creation and utility of standardized common data models, such as OMOP [44], allow for more wide-spread replication analyses but it does not overcome the limitations of joint data access. As such, it is imperative that alternative strategies emerge for learning from multiple EHR data sources that go beyond the common discovery-replication framework. Federated learning might be the tool to enable large-scale representative ML of EHR data and we discuss many studies which demonstrate this fact below.

Federated learning is a viable method to connect EHR data from medical institutions, allowing them to share their experiences, and not their data, with a guarantee of privacy [9, 25, 34, 45, 65, 82]. In these scenarios, the performance of ML model will be significantly improved by the iterative improvements of learning from large and diverse medical data sets. There have been some tasks were studied in federated learning setting in healthcare, e.g., patient similarity learning [62], patient representation learning, phenotyping [55, 67], and predictive modeling [10, 45, 90]. Specifically, Lee et al. [62] presented a privacy-preserving platform in a federated setting for patient similarity learning across institutions. Their model can find similar patients from one hospital to another without sharing patient-level information. Kim et al. [55] used tensor factorization models to convert massive electronic health records into meaningful phenotypes for data analysis in federated learning setting. Liu et al. [67] conducted both patient representation learning and obesity comorbidity phenotyping in a federated manner and got good results. Vepakomma et al. [103] built several configurations upon a distributed deep learning method called SplitNN [36] to facilitate the health entities collaboratively training deep learning models without sharing sensitive raw data or model details. Silva et al. [93] illustrated their federated learning framework by investigating brain structural relationships across diseases and clinical cohorts. Huang et al. [45] sought to tackle the challenge of non-IID ICU patient data by clustering patients into clinically meaningful communities that

captured similar diagnoses and geological locations and simultaneously training one model per community.

Federated learning has also enabled predictive modeling based on diverse sources, which can provide clinicians with additional insights into the risks and benefits of treating patients earlier [9, 10, 90]. Brisimi et al. [10] aimed to predict future hospitalizations for patients with heart-related diseases using EHR data spread among various data sources/agents by solving the l_1-regularized sparse Support Vector Machine classifier in federated learning environment. Owkin is using federated learning to predict patients' resistance to certain treatment and drugs, as well as their survival rates for certain diseases [99]. Boughorbel et al. [9] proposed a federated uncertainty-aware learning algorithm for the prediction of preterm birth from distributed EHR, where the contribution of models with high uncertainty in the aggregation model is reduced. Pfohl et al. [80] considered the prediction of prolonged length of stay and in-hospital mortality across thirty-one hospitals in the eICU Collaborative Research Database. Sharma et al. [90] tested a privacy preserving framework for the task of in-hospital mortality prediction among patients admitted to the intensive care unit (ICU). Their results show that training the model in the federated learning framework leads to comparable performance to the traditional centralized learning setting. Summary of these work is listed in Table 1.

3.2 Others

An important application of federated learning is for natural language processing (NLP) tasks. When Google first proposed federated learning concept in 2016, the application scenario is Gboard—a virtual keyboard of Google for touchscreen mobile devices with support for more than 600 language varieties [39, 70]. Indeed, as users increasingly turn to mobile devices, fast mobile input methods with auto-correction, word completion, and next-word prediction features are becoming more and more important. For these NLP tasks, especially next-word prediction, typed text in mobile apps is usually better than the data from scanned books or speech-to-text in terms of aiding typing on a mobile keyboard. However, these language data often contain sensitive information, e.g., passwords, search queries, or text messages with personal information. Therefore, federated learning has a promising application in NLP like virtual keyboard prediction [7, 39, 70].

Other applications include smart retail [112] and finance [54]. Specifically, smart retail aims to use machine learning technology to provide personalized services to customers based on data like user purchasing power and product characteristics for product recommendation and sales services. In terms of financial applications, Tencent's WeBank leverages federated learning technologies for credit risk management, where several Banks could jointly generate a comprehensive credit score for a customer without sharing his or her data [109]. With the growth and development of federated learning, there are many companies or research teams that have carried out various tools oriented to scientific research and product development. Popular ones are listed in Table 2.

Table 1 Summary of recent work on federated learning for healthcare

Problem	ML method	No. of clients	Data
Patient similarity learning [62]	Hashing	3	MIMIC-III [50]
Patient similarity learning [108]	Hashing	20	MIMIC-III
Phenotyping [55]	TF	1–5	MIMIC-III, UCSD [104]
Phenotyping [67]	NLP	10	MIMIC-III
Representation learning [93]	PCA	10–100	ADNI, UK Biobank, PPMI, MIRIAD
Mortality prediction [45]	Autoencoder	5–50	eICU Collaborative Research Database [81]
Hospitalization prediction [10]	SVM	5, 10	Boston Medical Center
Preterm-birth prediction [9]	RNN	50	Cerner Health Facts
Mortality prediction [80]	LR, NN	31	eICU Collaborative Research Database
Mortality prediction [90]	LR, MLP	2	MIMIC-III
Activity recognition [16]	CNN	5	UCI Smartphone [4]
Adverse drug reactions Prediction [19, 20]	SVM, MLP, LR	10	LCED, MIMIC
Arrhythmia detection [110]	NN	16, 32, 64	PhysioNet Dataset [21]
Disease prediction [33]	NN	5, 10	Pima Indians Diabetes Dataset [95], Cleveland Heart Disease Database [23]
Imaging data analysis	VAE	4	MNIST, Brain Imaging Data
Mortality prediction [101]	LRR, MLP, LASSO	5	Mount Sinai COVID-19 Dataset

TF tensor factorization, *MLP* multi-layer perceptron, *VAE* variational autoencoder, *LCED* Limited MarketScan Explorys Claims-EMR Data. https://www.ibm.com/downloads/cas/6KNYVVQ2

4 Conclusions and Open Questions

In this survey, we review the current progress on federated learning including, but not limited to healthcare field. We summarize the general solutions to the various challenges in federated learning and hope to provide a useful resource for researchers to refer. Besides the summarized general issues in federated learning setting, we list some probably encountered directions or open questions when federated learning is applied in healthcare area in the following.

– **Data Quality**. Federated learning has the potential to connect all the isolated medical institutions, hospitals, or devices to make them share their experiences with privacy guarantee. However, most health systems suffer from data clutter and efficiency problems. The quality of data collected from multiple sources is uneven and there is no uniform data standard. The analyzed results are apparently worthless when dirty data are accidentally used as samples. The ability to strategically leverage medical data is critical. Therefore, how to clean, correct, and complete data and accordingly ensure data quality is a key to improve the

Table 2 Popular tools for federated learning research

Project name	Developer	Description
PySyft [87]	OpenMined	It decouples private data from model training using federated learning, DP, and MPC within PyTorch. TensorFlow bindings are also available [77].
TFF [31]	Google	With TFF, TensorFlow provides users with a flexible and open framework through which they can simulate distributed computing locally.
FATE [3]	Webank	FATE support the Federated AI ecosystem, where a secure computing protocol is implemented based on homomorphic encryption and MPC.
Tensor/IO [24]	Dow et al.	Tensor/IO is a lightweight cross-platform library for on-device machine learning, bringing the power of TensorFlow and TensorFlow Lite to iOS, Android, and React native applications.

machine learning model weather we are dealing with federated learning scenario or not.

– **Incorporating Expert Knowledge**. In 2016, IBM introduced Watson for Oncology, a tool that uses the natural language processing system to summarize patients' electronic health records and search the powerful database behind it to advise doctors on treatments. Unfortunately, some oncologists say they trust their judgment more than Watson tells them what needs to be done.[1] Therefore, hopefully doctors will be involved in the training process. Since every data set collected here cannot be of high quality, so it will be very helpful if the standards of evidence-based machine are introduced, doctors will also see the diagnostic criteria of artificial intelligence. If wrong, doctors will give further guidance to artificial intelligence to improve the accuracy of machine learning model during training process."
– **Incentive Mechanisms**. With the internet of things and the variety of third party portals, a growing number of smartphone healthcare apps are compatible with wearable devices. In addition to data accumulated in hospitals or medical centers, another type of data that is of great value is coming from wearable devices not only to the researchers but more importantly for the owners. However, during federated model training process, the clients suffer from considerable overhead in communication and computation. Without well-designed incentives, self-interested mobile or other wearable devices will be reluctant to participate

[1] http://news.moore.ren/industry/158978.htm

in federal learning tasks, which will hinder the adoption of federated learning [53]. How to design an efficient incentive mechanism to attract devices with high-quality data to join federated learning is another important problem.

– **Personalization**. Wearable devices are more focus on public health, which means helping people who are already healthy to improve their health, such as helping them exercise, practice meditation, and improve their sleep quality. How to assist patients to carry out scientifically designed personalized health management, correct the functional pathological state by examining indicators, and interrupt the pathological change process are very important. Reasonable chronic disease management can avoid emergency visits and hospitalization and reduce the number of visits. Cost and labor savings. Although there are some general work about federated learning personalization [48, 94], for healthcare informatics, how to combining the medical domain knowledge and make the global model be personalized for every medical institutions or wearable devices is another open question.
– **Model Precision**. Federated tries to make isolated institutions or devices share their experiences, and the performance of machine learning model will be significantly improved by the formed large medical dataset. However, the prediction task is currently restricted and relatively simple. Medical treatment itself is a very professional and accurate field. Medical devices in hospitals have incomparable advantages over wearable devices. And the models of Doc.ai could predict the phenome collection of one's biometric data based on its selfie, such as height, weight, age, sex, and BMI.[2] How to improve the prediction model to predict future health conditions is definitely worth exploring.

Funding The work is supported by ONR N00014-18-1-2585 and NSF 1750326. FW would also like to acknowledge the support from Amazon AWS Machine Learning Research Award and Google Faculty Research Award.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Abadi M, Chu A, Goodfellow I, McMahan HB, Mironov I, Talwar K, Zhang L (2016) Deep learning with differential privacy. In: Proceedings of the 2016 ACM SIGSAC conference on computer and communications security. ACM, pp 308–318
2. Agarwal N, Suresh AT, Yu FXX, Kumar S, McMahan B (2018) cpsgd: communication-efficient and differentially-private distributed sgd. In: Advances in neural information processing systems, pp 7564–7575
3. AI W (2019) Federated ai technology enabler. https://www.fedai.org/cn/
4. Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz JL (2012) Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In: International workshop on ambient assisted living. Springer, pp 216–223

[2] https://doc.ai/blog/do-you-know-how-valuable-your-medical-da/

5. Anh TT, Luong NC, Niyato D, Kim DI, Wang LC (2019) Efficient training management for mobile crowd-machine learning: A deep reinforcement learning approach. IEEE Wireless Communications Letters 8(5):1345–1348
6. Barcelos C, Gluz J, Vicari R (2011) An agent-based federated learning object search service. Interdisciplinary Journal of E-Learning and Learning Objects 7(1):37–54
7. Bonawitz K, Eichner H, Grieskamp W, Huba D, Ingerman A, Ivanov V, Kiddon C, Konecny J, Mazzocchi S, McMahan HB, et al. (2019) Towards federated learning at scale: System design. arXiv:1902.01046
8. Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan HB, Patel S, Ramage D, Segal A, Seth K (2017) Practical secure aggregation for privacy-preserving machine learning. In: Proceedings of the 2017 ACM SIGSAC conference on computer and communications security. ACM, pp 1175–1191
9. Boughorbel S, Jarray F, Venugopal N, Moosa S, Elhadi H, Makhlouf M (2019) Federated uncertainty-aware learning for distributed hospital ehr data. arXiv:1910.12191
10. Brisimi TS, Chen R, Mela T, Olshevsky A, Paschalidis IC, Shi W (2018) Federated learning of predictive models from federated electronic health records. Int J Med Inform 112:59–67
11. Bui TD, Nguyen CV, Swaroop S, Turner RE (2018) Partitioned variational inference: a unified framework encompassing federated and continual learning. arXiv:1811.11206
12. Caldas S, Konečny J, McMahan HB, Talwalkar A (2018) Expanding the reach of federated learning by reducing client resource requirements. arXiv:1812.07210
13. Caldas S, Wu P, Li T, Konečnỳ J, McMahan HB, Smith V, Talwalkar A (2018) Leaf: a benchmark for federated settings. arXiv:1812.01097
14. Chai D, Wang L, Chen K, Yang Q (2020) Secure federated matrix factorization. arXiv preprint arXiv:1906.05108
15. Chaudhuri K, Sarwate AD, Sinha K (2013) A near-optimal algorithm for differentially-private principal components. J Mach Learn Res 14(1):2905–2943
16. Chen Y, Qin X, Wang J, Yu C, Gao W (2020) Fedhealth: a federated transfer learning framework for wearable healthcare. IEEE Intelligent Systems
17. Chen Y, Sun X, Jin Y (2019) Communication-efficient federated deep learning with asynchronous model update and temporally weighted aggregation. arXiv:1903.07424
18. Cheng K, Fan T, Jin Y, Liu Y, Chen T, Yang Q (2019) Secureboost: a lossless federated learning framework. arXiv:1901.08755
19. Choudhury O, Gkoulalas-Divanis A, Salonidis T, Sylla I, Park Y, Hsu G, Das A (2019) Differential privacy-enabled federated learning for sensitive health data. arXiv:1910.02578
20. Choudhury O, Park Y, Salonidis T, Gkoulalas-Divanis A, Sylla I, et al. (2019) Predicting adverse drug reactions on distributed health data using federated learning. In: AMIA Annual symposium proceedings, vol 2019, p 313. American Medical Informatics Association
21. Clifford GD, Silva I, Moody B, Li Q, Kella D, Shahin A, Kooistra T, Perry D, Mark RG (2015) The physionet/computing in cardiology challenge 2015: reducing false arrhythmia alarms in the icu. In: 2015 Computing in Cardiology Conference (CinC). IEEE, pp 273–276
22. Corinzia L, Buhmann JM (2019) Variational federated multi-task learning. arXiv:1906.06268
23. Detrano R, Janosi A, Steinbrunn W, Pfisterer M, Schmid JJ, Sandhu S, Guppy KH, Lee S, Froelicher V (1989) International application of a new probability algorithm for the diagnosis of coronary artery disease. The American Journal of Cardiology 64(5):304–310
24. (2019). doc.ai: declarative, on-device machine learning for ios, android, and react native. https://github.com/doc-ai/tensorio
25. Duan R, Boland MR, Liu Z, Liu Y, Chang HH, Xu H, Chu H, Schmid CH, Forrest CB, Holmes JH, et al. (2020) Learning from electronic health records across multiple sites: a communication-efficient and privacy-preserving distributed algorithm. J Am Med Inform Assoc 27(3):376–385
26. Dwork C, Kenthapadi K, McSherry F, Mironov I, Naor M (2006) Our data, ourselves: Privacy via distributed noise generation. In: Annual international conference on the theory and applications of cryptographic techniques. Springer, pp 486–503
27. Dwork C, Rothblum GN, Vadhan S (2010) Boosting and differential privacy. In: 2010 IEEE 51st Annual symposium on foundations of computer science. IEEE, pp 51–60
28. Eichner H, Koren T, McMahan HB, Srebro N, Talwar K (2019) Semi-cyclic stochastic gradient descent. arXiv:1904.10120
29. Fontaine C, Galand F (2007) A survey of homomorphic encryption for nonspecialists. EURASIP J Inf Secur 2007:15

30. Glicksberg BS, Johnson KW, Dudley JT (2018) The next generation of precision medicine: observational studies, electronic health records, biobanks and continuous monitoring. Hum Mol Genet 27(R1):R56–R62
31. (2019). Google: Tensorflow federated. https://www.tensorflow.org/federated
32. Gostin LO (2001) National health information privacy: regulations under the health insurance portability and accountability act. JAMA 285(23):3015–3021
33. Grama M, Musat M, Muñoz-González L, Passerat-Palmbach J, Rueckert D, Alansary A (2020) Robust aggregation for adaptive privacy preserving federated learning in healthcare. arXiv:2009.08294
34. Gruendner J, Schwachhofer T, Sippl P, Wolf N, Erpenbeck M, Gulden C, Kapsner LA, Zierk J, Mate S, Stürzl M, et al. (2019) KETOS: Clinical decision support and machine learning as a service–A training and deployment platform based on Docker, OMOP-CDM, and FHIR Web Services. PloS one 14(10):1–16
35. Guha N, Talwalkar A, Smith V (2019) One-shot federated learning. arXiv:1902.11175
36. Gupta O, Raskar R (2018) Distributed learning of deep neural network over multiple agents. J Netw Comput Appl 116:1–8
37. Han S, Mao H, Dally WJ (2015) Deep compression: compressing deep neural networks with pruning, trained quantization and huffman coding. arXiv:1510.00149
38. Han Y, Zhang X (2019) Robust federated training via collaborative machine teaching using trusted instances. arXiv:1905.02941
39. Hard A, Rao K, Mathews R, Beaufays F, Augenstein S, Eichner H, Kiddon C, Ramage D (2018) Federated learning for mobile keyboard prediction. arXiv:1811.03604
40. Hardy S, Henecka W, Ivey-Law H, Nock R, Patrini G, Smith G, Thorne B (2017) Private federated learning on vertically partitioned data via entity resolution and additively homomorphic encryption. arXiv:1711.10677
41. He C, Tan C, Tang H, Qiu S, Liu J (2019) Central server free federated learning over single-sided trust social networks. arXiv:1910.04956
42. Hill P (1985) The Rationale for Learning Communities and Learning Community Models. ERIC
43. Hinton G, Vinyals O, Dean J (2015) Distilling the knowledge in a neural network. arXiv:1503.02531
44. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, Suchard MA, Park RW, Wong ICK, Rijnbeek PR, et al. (2015) Observational health data sciences and informatics (ohdsi): opportunities for observational researchers. Stud Health Technol Inform 216:574
45. Huang L, Liu D (2019) Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. arXiv:1903.09296
46. Ickin S, Vandikas K, Fiedler M (2019) Privacy preserving qoe modeling using collaborative learning. arXiv:1906.09248
47. Jensen PB, Jensen LJ, Brunak S (2012) Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet 13(6):395–405
48. Jiang Y, Konečnỳ J, Rush K, Kannan S (2019) Improving federated learning personalization via model agnostic meta learning. arXiv:1909.12488v1
49. Jin Y, Wei X, Liu Y, Yang Q (2020) A survey towards federated semi-supervised learning. arXiv:2002.11545
50. Johnson AE, Pollard TJ, Shen L, Li-wei HL, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG (2016) Mimic-iii, a freely accessible critical care database. Sci Data 3:160035
51. Kairouz P, McMahan HB, Avent B, Bellet A, Bennis M, Bhagoji AN, Bonawitz K, Charles Z, Cormode G, Cummings R, et al. (2019) Advances and open problems in federated learning. arXiv:1912.04977
52. Kamp M, Adilova L, Sicking J, Hüger F., Schlicht P, Wirtz T, Wrobel S (2018) Efficient decentralized deep learning by dynamic model averaging. In: Joint European conference on machine learning and knowledge discovery in databases. Springer, pp 393–409
53. Kang J, Xiong Z, Niyato D, Yu H, Liang YC, Kim DI (2019) Incentive design for efficient federated learning in mobile networks: A contract theory approach. arXiv:1905.07479
54. Kawa D, Punyani S, Nayak P, Karkera A, Jyotinagar V (2019) Credit risk assessment from combined bank records using federated learning. International Research Journal of Engineering and Technology (IRJET) 6(4):1355–1358

55. Kim Y, Sun J, Yu H, Jiang X (2017) Federated tensor factorization for computational phenotyping. In: Proceedings of the 23rd ACM SIGKDD International conference on knowledge discovery and data mining. ACM, pp 887–895
56. Konečnỳ J, McMahan B, Ramage D (2015) Federated optimization: distributed optimization beyond the datacenter. arXiv:1511.03575
57. Konečnỳ J, McMahan HB, Ramage D, Richtárik P (2016) Federated optimization: distributed machine learning for on-device intelligence. arXiv:1610.02527
58. Konečnỳ J, McMahan HB, Yu FX, Richtárik P, Suresh AT, Bacon D (2016) Federated learning: strategies for improving communication efficiency. arXiv:1610.05492
59. Kulkarni V, Kulkarni M, Pant A (2020) Survey of personalization techniques for federated learning. arXiv:2003.08673
60. Lalitha A, Kilinc OC, Javidi T, Koushanfar F (2019) Peer-to-peer federated learning on graphs. arXiv:1901.11173
61. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
62. Lee J, Sun J, Wang F, Wang S, Jun CH, Jiang X (2018) Privacy-preserving patient similarity learning in a federated environment: development and analysis. JMIR Medical Informatics 6(2):e20
63. Li T, Sahu AK, Zaheer M, Sanjabi M, Talwalkar A, Smith1 V (2019) Federated optimization for heterogeneous networks. arXiv:1812.06127
64. Li T, Sanjabi M, Smith V (2019) Fair resource allocation in federated learning. arXiv:1905.10497
65. Li Z, Roberts K, Jiang X, Long Q (2019) Distributed learning from multiple ehr databases: Contextual embedding models for medical events. J Biomed Inform 92:103138
66. Lim WYB, Luong NC, Hoang DT, Jiao Y, Liang YC, Yang Q, Niyato D, Miao C (2019) Federated learning in mobile edge networks: a comprehensive survey. arXiv:1909.11875
67. Liu D, Dligach D, Miller T (2019) Two-stage federated phenotyping and patient representation learning. arXiv:1908.05596
68. Lyu L, Yu H, Yang Q (2020) Threats to federated learning: a survey. arXiv:2003.02133
69. McMahan B, Moore E, Ramage D, Hampson S, Arcas BA (2017) Communication-efficient learning of deep networks from decentralized data. In: Artificial intelligence and statistics, pp 1273–1282
70. McMahan HB, Ramage D, Talwar K, Zhang L (2017) Learning differentially private recurrent language models. arXiv:1710.06963
71. Min X, Yu B, Wang F (2019) Predictive modeling of the hospital readmission risk from patients' claims data using machine learning: A case study on copd. Sci Rep 9(1):2362
72. Miotto R, Wang F, Wang S, Jiang X, Dudley JT (2018) Deep learning for healthcare: review, opportunities and challenges. Brief Bioinformatics 19(6):1236–1246
73. Mohri M, Sivek G, Suresh AT (2019) Agnostic federated learning. In: Chaudhuri K, Salakhutdinov R (eds) Proceedings of the 36th International conference on machine learning, proceedings of machine learning research, vol 97. PMLR, Long Beach, pp 4615–4625
74. Mukherjee R, Jaffe H (2005) System and method for dynamic context-sensitive federated search of multiple information repositories. US Patent App. 10/743,196
75. Nishio T, Yonetani R (2018) Client selection for federated learning with heterogeneous resources in mobile edge. arXiv:1804.08333
76. Obermeyer Z, Powers B, Vogeli C, Mullainathan S (2019) Dissecting racial bias in an algorithm used to manage the health of populations. Science 366(6464):447–453
77. OpenMined: Pysyft-tensorflow. https://github.com/OpenMined/PySyft-TensorFlow (2019)
78. Pathak M, Rane S, Raj B (2010) Multiparty differential privacy via aggregation of locally trained classifiers. In: Advances in neural information processing systems, pp 1876–1884
79. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, Balasubramanian V, Russo AM, Rajmane A, Cheung L, et al. (2019) Large-scale assessment of a smartwatch to identify atrial fibrillation. N Engl J Med 381(20):1909–1917
80. Pfohl SR, Dai AM, Heller K (2019) Federated and differentially private learning for electronic health records. arXiv:1911.05861
81. Pollard TJ, Johnson AE, Raffa JD, Celi LA, Mark RG, Badawi O (2018) The eicu collaborative research database, a freely available multi-center database for critical care research. Sci Data 5:180178
82. Raja PV, Sivasankar E (2014) Modern framework for distributed healthcare data analytics based on hadoop. In: Information and communication technology-EurAsia conference. Springer, pp 348–355

83. Rehak D, Dodds P, Lannom L (2005) A model and infrastructure for federated learning content repositories. In: Interoperability of web-based educational systems workshop, vol 143. Citeseer
84. Ren J, Yu G, Ding G (2019) Accelerating dnn training in wireless federated edge learning system. arXiv:1905.09712
85. Roy AG, Siddiqui S, Pölsterl S, Navab N, Wachinger C (2019) Braintorrent: a peer-to-peer environment for decentralized federated learning. arXiv:1905.06731
86. Rubinstein BI, Bartlett PL, Huang L, Taft N (2009) Learning in a large function space: privacy-preserving mechanisms for svm learning. arXiv:0911.5708
87. Ryffel T, Trask A, Dahl M, Wagner B, Mancuso J, Rueckert D, Passerat-Palmbach J (2018) A generic framework for privacy preserving deep learning. arXiv:1811.04017
88. Samarakoon S, Bennis M, Saad W, Debbah M (2018) Federated learning for ultra-reliable low-latency v2v communications. In: 2018 IEEE Global Communications Conference (GLOBECOM). IEEE, pp 1–7
89. Sattler F, Wiedemann S, Müller K. R., Samek W (2019) Robust and communication-efficient federated learning from non-iid data. arXiv:1903.02891
90. Sharma P, Shamout FE, Clifton DA (2019) Preserving patient privacy while training a predictive model of in-hospital mortality. arXiv:1912.00354
91. Shayan M, Fung C, Yoon CJ, Beschastnikh I (2018) Biscotti: a ledger for private and secure peer-to-peer machine learning. arXiv:1811.09904
92. Shokri R, Shmatikov V (2015) Privacy-preserving deep learning. In: Proceedings of the 22nd ACM SIGSAC conference on computer and communications security. ACM, pp 1310–1321
93. Silva S, Gutman B, Romero E, Thompson PM, Altmann A, Lorenzi M (2018) Federated learning in distributed medical databases: meta-analysis of large-scale subcortical brain data. arXiv:1810.08553
94. Sim KC, Zadrazil P, Beaufays F (2019) An investigation into on-device personalization of end-to-end automatic speech recognition models. arXiv:1909.06678
95. Smith JW, Everhart J, Dickson W, Knowler W, Johannes R (1988) Using the adap learning algorithm to forecast the onset of diabetes mellitus. In: Proceedings of the annual symposium on computer application in medical care, p 261. American medical informatics association
96. Smith V, Chiang CK, Sanjabi M, Talwalkar AS (2017) Federated multi-task learning. In: Advances in neural information processing systems, pp 4424–4434
97. Such JM, Criado N (2018) Multiparty privacy in social media. Commun ACM 61(8):74–81
98. Thomas K, Grier C (2010) Nicol, D.M.: unfriendly: multi-party privacy risks in social networks. In: International symposium on privacy enhancing technologies symposium. Springer, pp 236–252
99. Tramel E (2019) Federated learning: rewards & challenges of distributed private ml. Accessed May 28, 2019
100. Truex S, Baracaldo N, Anwar A, Steinke T, Ludwig H, Zhang R (2018) A hybrid approach to privacy-preserving federated learning. arXiv:1812.03224
101. Vaid A, Jaladanki SK, Xu J, Teng S, Kumar A, Lee S, Somani S, Paranjpe I, De Freitas JK, Wanyan T, et al. (2020) Federated learning of electronic health records improves mortality prediction in patients hospitalized with covid-19 medRxiv
102. Van Hasselt H, Guez A, Silver D (2016) Deep reinforcement learning with double q-learning. In: Thirtieth AAAI conference on artificial intelligence
103. Vepakomma P, Gupta O, Swedish T, Raskar R (2018) Split learning for health: distributed deep learning without sharing raw patient data. arXiv:1812.00564
104. Wah C, Branson S, Welinder P, Perona P, Belongie S (2011) The caltech-ucsd birds-200-2011 dataset
105. Wang F, Preininger A (2019) Ai in health: state of the art, challenges, and future directions. Yearb Med Inform 28(01):016–026
106. Wang X, Han Y, Wang C, Zhao Q, Chen X, Chen M (2018) In-edge ai: Intelligentizing mobile edge computing, caching and communication by federated learning. arXiv:1809.07857
107. Xu J, Wang F (2019) Federated learning for healthcare informatics. arXiv:1911.06270
108. Xu J, Xu Z, Walker P, Wang F (2020) Federated patient hashing. In: AAAI, pp 6486–6493
109. Yang Q, Liu Y, Chen T, Tong Y (2019) Federated machine learning: concept and applications. ACM Trans Intell Syst Technol 10(2):12:1–12:19. https://doi.org/10.1145/3298981
110. Yuan B, Ge S, Xing W (2020) A federated learning framework for healthcare iot devices. arXiv:2005.05083
111. Zhao Y, Li M, Lai L, Suda N, Civin D, Chandra V (2018) Federated learning with non-iid data. arXiv:1806.00582

112. Zhao Y, Zhao J, Jiang L, Tan R, Niyato D (2019) Mobile edge computing, blockchain and reputation-based crowdsourcing iot federated learning: a secure, decentralized and privacy-preserving system. arXiv:1906.10893
113. Zhu H, Jin Y (2019) Multi-objective evolutionary federated learning. IEEE transactions on neural networks and learning systems
114. Zhuo HH, Feng W, Xu Q, Yang Q, Lin Y (2019) Federated reinforcement learning. arXiv:1901.08277