Federated Learning for Healthcare: Systematic Review and Architecture Proposal

RODOLFO STOFFEL ANTUNES and CRISTIANO ANDRÉ DA COSTA, Universidade do Vale do Rio dos Sinos, São Leopoldo, Brazil
ARNE KÜDERLE, IMRANA ABDULLAHI YARI, and BJÖRN ESKOFIER,
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

The use of **machine learning (ML)** with **electronic health records (EHR)** is growing in popularity as a means to extract knowledge that can improve the decision-making process in healthcare. Such methods require training of high-quality learning models based on diverse and comprehensive datasets, which are hard to obtain due to the sensitive nature of medical data from patients. In this context, **federated learning (FL)** is a methodology that enables the distributed training of machine learning models with remotely hosted datasets without the need to accumulate data and, therefore, compromise it. FL is a promising solution to improve ML-based systems, better aligning them to regulatory requirements, improving trustworthiness and data sovereignty. However, many open questions must be addressed before the use of FL becomes widespread. This article aims at presenting a systematic literature review on current research about FL in the context of EHR data for healthcare applications. Our analysis highlights the main research topics, proposed solutions, case studies, and respective ML methods. Furthermore, the article discusses a general architecture for FL applied to healthcare data based on the main insights obtained from the literature review. The collected literature corpus indicates that there is extensive research on the privacy and confidentiality aspects of training data and model sharing, which is expected given the sensitive nature of medical data. Studies also explore improvements to the aggregation mechanisms required to generate the learning model from distributed contributions and case studies with different types of medical data.

CCS Concepts: • **Applied computing** → **Health informatics**; • **Computer systems organization** → **Distributed architectures**; • **Computing methodologies** → **Machine learning approaches;**

Additional Key Words and Phrases: Electronic health records, federated learning, systematic review

ACM Reference format:
Rodolfo Stoffel Antunes, Cristiano André da Costa, Arne Küderle, Imrana Abdullahi Yari, and Björn Eskofier. 2022. Federated Learning for Healthcare: Systematic Review and Architecture Proposal. *ACM Trans. Intell. Syst. Technol.* 13, 4, Article 54 (May 2022), 23 pages.
https://doi.org/10.1145/3501813

This work is supported by: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) financial code 0001; Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) ARD 04/2019 grant 19/2551-0001340-0; and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant 309537/2020-7.
Authors' addresses: R. S. Antunes and C. A. da Costa, Universidade do Vale do Rio dos Sinos, Av. Unisinos 950, CEP 93022-750, São Leopoldo, RS, Brazil; emails: {rsantunes, cac}@unisinos.br; A. Küderle, I. A. Yari, and B. Eskofier, Friedrich-Alexander-Universität, Carl-Thiersch-Straße 2b, 91052 Erlangen, Bayern, Germany; emails: {arne.kuederle, imrana.yari.abdullahi, bjoern.eskofier}@fau.de.

2157-6904/2022/05-ART54 $15.00
https://doi.org/10.1145/3501813

1 INTRODUCTION

The pervasiveness of computational systems in healthcare environments led to the widespread availability of electronic health data [1]. These datasets enable the use of advanced data analysis techniques based on **machine learning (ML)** to extract vast amounts of knowledge that can improve various aspects of healthcare [2]. Ultimately, the analysis of electronic health data using ML techniques can result in treatments and procedures with lower risks and better outcomes for patients, thus increasing the quality of care [3].

Most ML methods applied to healthcare focus on the analysis of **Electronic Health Records (EHR)** data, which contains the individual medical information from a person, possibly including appointments, diagnoses, treatments, and exams such as medical imaging [4]. EHR emerged from the natural adoption of information technology by medical institutions and became a relevant source of data for many analytical methods aimed at extracting knowledge to aid medical practices [5]. In particular, recent work shows the potential of deep learning methods applied to EHR data to extract knowledge that can directly improve the outcomes and quality of care [6].

ML-based methods require a learning model trained to identify patterns and extract the intended knowledge from raw data. This training step is crucial to the accuracy of the resulting analysis method. It must be conducted with a large dataset with a diverse number of samples to guarantee the ML model quality [7]. The training's results directly depend on the number of data items, the diversity of the samples, and the quality of dataset annotation regarding the expected classification. Acquiring or producing such a dataset often is a time-consuming and expensive task. This procedure involves several parties collecting the data, transferring it to a central data repository, and fusing it to build a model. In turn, data owners may be unaware of these procedures and the future use cases of the ML model. For that reason, in the particular case of healthcare, the traditional ML processes may violate laws such as the **General Data Protection Regulation (GDPR)** of the European Union, the **California Consumer Privacy Act (CCPA)**, and **Health Insurance Portability and Accountability Act (HIPAA)** [8]. Such difficulties and restrictions on medical data sharing are the major hindering factor in the development of advanced ML techniques for healthcare [9].

Recent work proposes the concept of **federated learning (FL)** [10] to tackle these issues. In summary, FL enables the training of ML models locally (at the location of the data) and only shares the resulting model, which is not reverse-engineerable, with the requesting party. Therefore, FL avoids the need to share the private datasets and sensitive data to others, preventing exposition to entities conducting studies and enabling data usage for broader purposes [11]. A central entity manages the learning process and distributes the training algorithm to each participating data holder. Each participant generates a *local model* trained with their private data and shares the resulting parameters with the central entity. Finally, the central entity employs an aggregation algorithm to combine the parameters of all local models into a single global model.

Applications described in the literature successfully employ the methodology described above. For example, Hard et al. [12] train the prediction model of a mobile keyboard using private data stored in smartphones. Similar studies demonstrated the viability and proposed solutions enabling FL in a wide variety of contexts. In particular, FL provides a solution to train high-quality models for ML applied to healthcare [13]. As a result, a group of studies aims to evaluate the state-of-the-art on this subject. Xu et al. [14] briefly discuss the applications of FL on healthcare applications. Rieke et al. [15] explore various technical and social aspects related to FL in healthcare applied to healthcare. Zerka et al. [16] survey the FL methods with the specific focus of their advantages and drawbacks to the privacy and security of medical data. Li et al. [17] present a general overview of FL applications but does not provide an in-depth analysis of challenges related to medical data.

Mothukuri et al. [18] focus on the security aspects of FL architectures and strategies to mitigate threats. Finally, Zhang et al. [19] survey FL research from the perspective of architectural challenges and future directions. These studies provide a general overview of FL, its application to healthcare, and the implications regarding technical and social perspectives. However, none provide an in-depth exploration of current literature in light of FL applications for healthcare, classifying it according to the employed methodologies and technical challenges. Furthermore, none of the studies leverages the researched challenges and solutions to propose a general architecture to enable FL for studies that require datasets owned by healthcare institutions.

This article aims at providing a comprehensive review of current studies related to the application of FL to the analysis of EHR data with ML-enabled methods. The study follows a systematic literature review methodology [20] with a well-defined search methodology to select a corpus of 44 recent articles for analysis. In summary, our review provides the following contributions:

- A comprehensive literature review of current work on FL applied to EHR data, including main research issues and proposed solutions, employed ML techniques, and case studies.
- A discussion about a general architecture to facilitate the use of FL with ML-enabled applications for healthcare, including the concept of federated data analysis.

The remainder of this paper is organized as follows: Section 2 describes the methodology used to search and select the articles that comprise the literature corpus. Section 3 presents the results of our literature analysis, focusing on answering a set of well-defined research questions. Section 4 overviews an architecture to facilitate the use of FL and analysis with distributed EHR data. Finally, Section 5 provides some final remarks and future perspectives related to our work.

2 METHODOLOGY

The methodology adopted in this survey follows the principles of a systematic review [20]. It comprises a set of well-defined and documented steps that favor the reproducibility of obtained results. In summary, these steps are: *(i)* definition of research questions; *(ii)* selection of keyword and literature databases; *(iii)* article filtering and selection; and *(iv)* result analysis and discussion. The remainder of this section describes in more detail the adopted methodology.

The first step is to define the research questions to answer with the literature review. This work comprises one **main question (MQ)** and four **specific questions (SQ)**, as listed below:

MQ What are the current applications of FL in the context of EHR analysis?
SQ1 What are the main research questions and respective solutions discussed in the literature?
SQ2 How do we classify the current work on FL applied to EHR?
SQ3 What are issues that remain as open research questions on FL applied to EHR?
SQ4 What is the design of an FL architecture that leverages EHR data?

Using the research questions, we can define the keywords to extract the raw corpus from literature databases. The keywords comprise the most relevant terms in the main research question, including common synonyms. Regarding the literature databases, we employ those most relevant to the fields of computer science and medicine. Table 1 lists the databases used in this research and the employed keywords. These parameters resulted in an initial raw literature corpus of 200 articles.

The next phase of the methodology is to analyze and filter the raw corpus, aiming to obtain the final corpus for the proposed literature review. The filtering process comprises a series of steps that aim to eliminate results that do not add to the quality of the resulting literature review. It comprises the following criteria:

Table 1. Databases and Keywords Used to Extract the Raw Literature Corpus

Literature Databases	
Name	**Address**
ACM Digital Library	https://dl.acm.org
IEEE Xplore	https://plore.ieee.org
SpringerLink	https://link.springer.com
ScienceDirect	https://www.sciencedirect.com
PubMed	https://www.ncbi.nlm.nih.gov/pubmed
arXiv	https://arxiv.org
JMIR	https://jmirpublications.com

Keywords
"federated learning" ∧ ("electronic health records" ∨ "electronic medical records" ∨ "EHR" ∨ "EMR")

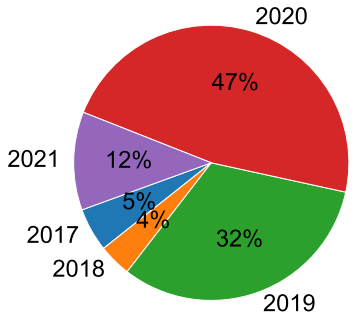

Fig. 1. Distribution of articles by publication year.

(1) **Publication date:** the corpus should include articles published in the last five years (2016 to 2021) to limit the review to the most recent work in the area.
(2) **Title and abstract review:** provides an initial screening to remove those not directly related to FL applied to EHR.
(3) **Duplicate removal:** combines articles from different databases and remove any existing duplicates.
(4) **Quality assessment:** remove articles that do not contain all elements expected in a complete research work.

The application of the filtering criteria results in a set of 67 articles. These comprise the final literature corpus used in this review. An initial analysis of the articles shows that FL applied to EHR is a recent research area. Figure 1 illustrates this aspect with the distribution of articles by publication year. In particular, we observe that the final corpus contains 47% of articles published in 2020 and 12% in the first quarter of 2021.

The final step of the systematic review methodology comprises the analysis and discussion of the literary corpus. The next section presents the results from this step, considering the research questions proposed for this review.

3 LITERATURE ANALYSIS

This section presents the analysis of the literature corpus collected with the described methodology. First, Section 3.1 explores the evolution of studies on federated learning applied to healthcare, including the main research issues and proposed solutions. Next, Section 3.2 summarizes the

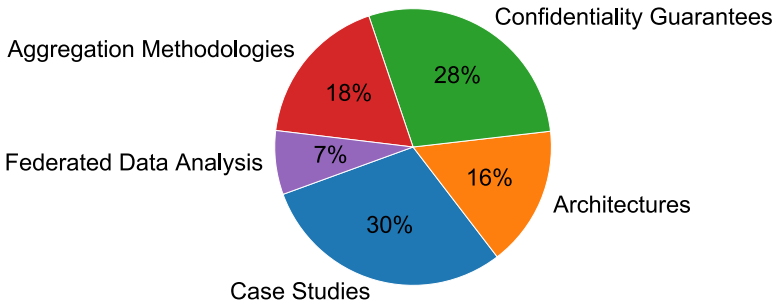

Fig. 2. Distribution by topics of articles in the literature corpus.

observed ML methods and case studies observed in the literature. Finally, Section 3.3 presents a taxonomy to classify the observed literature and discusses the main insights obtained from our analysis.

3.1 State-of-the-Art Evolution

The literature corpus shows that FL applied to EHR data analysis is a very recent research field, with most publications from the last two years. Earlier publications focus on architecture definitions and experimental case studies to evaluate the accuracy of learning models trained with FL concepts. Some earlier studies also focus on evaluating the privacy guarantees of the FL process in light of the stringent security requirements from medical data analysis. In turn, recent work added discussions on specific challenges that hinder the accuracy and confidentiality of learning models generated with FL. In particular, they explore the aggregation and cryptographic mechanisms employed in FL architectures.

Our analysis classifies the literature corpus in five general topics: *(i)* **case studies** of FL on medical data; *(ii)* **architectures** for FL in medical applications; *(iii)* **confidentiality guarantees** for training data; *(iv)* **aggregation methodologies** for global model generation; and *(v)* **federated data analysis**. Figure 2 illustrates the distribution by topic of the articles in the literature corpus. The remainder of this section explores the studies that focus on each of these topics.

3.1.1 Case Studies. This group encompasses some of the first work on FL applied to EHR data. These studies aim to apply the concepts presented by Yang et al. [21] to explore the viability of FL as a tool for training learning models over private medical data, including EHR. In this group of early studies, Jochems et al. [22] employ distributed learning to train a Bayesian network and evaluate the long-term evolution of lung cancer patients. Kim et al. [23] propose a federated tensor factorization method that iteratively trains and aggregates a learning model to extract phenotypes from EHR datasets. Sheller et al. [24] explore a distributed architecture to train a deep neural network to segment MRI scans from patients with brain tumors. Liu et al. [25] develop a distributed two-layer neural network for mortality prediction based on EHR data, in which the bottom layer is private to each participant and the top layer is public and trained using the aggregation of anonymous parameters from local models.

Regarding more recent studies, Liu et al. [26] employ a deep averaging network to extract medical datasets from plain text health records. The result is used in a distributed architecture based on support vector machines to train models to predict diagnoses based on EHR data. Silva et al. [27] explore particular mathematical characteristics of learning models for neurological MRI data analysis and propose a federated training architecture specific for this scenario. Their goal

is to train high-accuracy models for neurological disease diagnosis. Jiang et al. [28] propose an FL architecture focused on models for topic analysis on distributed datasets. This type of analysis focuses on identifying the general subjects from a given text and can be applied to analyze medical data in textual format. Sharma et al. [29] compares a learning model trained with a federated architecture against one with a centralized method. Their results indicate that FL can achieve an accuracy level similar to that of centralized methods, and the additional privacy is advantageous to healthcare applications. Ribero et al. [30] employ k-means clustering to learn prototypes from local dataset parameters that can be shared without compromising data confidentiality. These prototypes can be aggregated into a global learning model that drives recommendation systems on different applications, including medical data. Deist et al. [31] evaluate a case study of FL for generating regression models for two-year survival prediction of lung cancer patients. The case study explores an architecture deployed at five institutions that holds a total of 20,000 patient records. Szatmari et al. [32] propose a standard FL architecture to train learning models for the personalization of hearing aids. Sheller et al. [33] apply FL to train a model to identify cancer-affected brain tissue using as a case study the BraTS 2017 dataset. Vaid et al. [34] evaluate the results from traditional and FL training using as case study COVID-19 mortality prediction using EHR data including demographics, medical history, and vital signs. Lee et al. [35] explore the case study of ultrasound image analysis for tumor identification to compare FL and traditional training architectures. Liu et al. [36] employ FL to predict the mental state of patients based on physiological and motion data collected with medical IoT devices.

Another group of recent studies focuses on FL applied to data analysis from resource-constrained **Internet of health things (IoHT)** [37] devices. Guo et al. [38] propose a solution based on edge computing to enable FL in scenarios with resource-constrained devices. Chen et al. [39] explore transfer learning to improve the accuracy of global models trained to analyze IoHT data. Ju et al. [40] explore transfer learning to improve the accuracy of models to analyze **EEG (electroencephalogram)** data to enable brain-computer interfaces. Zhang et al. [41] propose an FL tailored to train models and analyze **ECG (electrocardiogram)** data to diagnose arrhythmia using IoHT devices. Can and Ersoy [42] also explore FL for ECG data analysis but focuses on detecting stress. Lim et al. [43] explore an incentive mechanism to guarantee that FL peers remain active in long training sessions for learning models for IoHT data.

The described work focuses on employing FL mechanisms to different ML techniques for medical data analysis. While deep neural networks stand out as a common technique for multiple studies, there is interest in exploring other techniques, for example, Bayesian networks and k-means clustering, which are well known for medical applications. Furthermore, these studies do not present major proposals for specific FL challenges, such as model aggregation or data confidentiality and integrity. Their main goal is to compare the precision of a model trained in a distributed architecture against one that employs traditional, centralized training. Furthermore, some qualitatively explore the trade-off between the loss of precision that stems from distributed training and the inherent advantages of FL (e.g., access to more varied datasets).

In general, these studies show that learning models generated with FL provide equal or better accuracy than those trained with centralized methodologies. These results stem from the broader information available from the datasets distributed in multiple institutions, which would not be available in circumstances where patient anonymity cannot be guaranteed. Studies also indicate that FL provides similar benefits to distributed ML, such as load distribution for model training. However, studies also show that the characteristics of individual datasets can directly impact training results, mainly if there is a high deviation in class distribution among datasets. In those situations, the FL method also needs to account for such class disparities during training.

3.1.2 Architectures. The second group of studies turn their attention to the proposal of complete architectures for FL applied to medical data. They have a greater focus on the combination of techniques and the overall distributed architecture required for FL. General studies in this group include the work from Brisimi et al. [44], which explores the design of a distributed architecture to solve large-scale sparse **support vector machine (SVM)** models. The resulting architecture enables the use of SVM in FL environments for medical data analysis. Lu et al. [45] propose a completely distributed version of the stochastic gradient descent algorithm, which enables an FL environment that does not require a central entity for model aggregation (peer-to-peer or serverless FL framework). Their study focuses on simulated studies with mathematical models that verify the correctness of their aggregation algorithm. Both studies explore a distributed version of a specific FL component and then describe a complete architecture around it. Also, these proposals aim to process historical EHR data to predict the diagnosis of specific cardiac and mental conditions.

Some studies focus on adapting well-known distributed architecture designs and explore their advantages in the area of FL. Roy et al. [46] employ the concept of BitTorrent swarms to distribute local model updates to all participants on an FL architecture. Each client individually maintains the global model according to the received local updates. Sanyal et al. [47] propose the use of a fog infrastructure to enable FL on constrained IoT medical devices. A fog infrastructure encompasses computational resources available in a hospital's local network, which can be used to manage complex tasks that otherwise would require cloud computing. Such an infrastructure enables the resources of constrained devices to be dedicated to data collection and simple learning tasks. Xue et al. [48] explore software-defined networks and mobile edge computing as architectures to deploy an FL system for e-health treatment recommendation based on IoHT devices. Polap et al. [49] propose a multi-agent infrastructure based on FL to train models for medical diagnosis using distributed databases. Finally, Hao et al. [50] propose a method to partition neural network models for distributed training. The method enables clients to employ a cloud platform for part of the training process, reducing the required computational power. The authors also propose a data perturbation method to guarantee data privacy when sharing partial models. These studies explore how different architectures can be used to distribute the resources required by the FL process, including computational resources and local models. They focus on two general types of applications. The first is the time-series analysis of medical data, particularly for sepsis outcome prediction or activity monitoring of patients. The second is medical image segmentation and classification of diagnostic prediction of different types of tumors.

Given the popularity of blockchain as distributed architectures for data authenticity, a group of studies also explore their applicability to FL. Shae and Tsai [51] propose to use a blockchain to store anonymous public models that can be freely aggregated by any member of the architecture that requires a learning model. Lugan et al. [52] further explore the properties of blockchains and aims to guarantee the authenticity of local learning models provided by individual clients. Clients also participate in a verification process that guarantees that new contributions improve the accuracy of the global learning model before their publication. Marulli et al. [53] aim to increase the confidentiality guarantees of learning models maintained in the blockchain with the use of homomorphic encryption. Rahman et al. [54] employ a blockchain to share partial training data so clients can download and train advanced models. Partial models employ differential privacy to guarantee privacy. All the studies employ the blockchain to authenticate and audit contributions from peers to the global parameter pool used to generate the global learning model. These studies focus on a varied number of case studies. Particular applications include image analysis for tissue recognition or disease diagnostics. In particular, Rahman et al. [54] explore a wide range of learning model applications tailored for COVID-19 management, such as infection contact management and outbreak management.

In summary, the studies above explore different distributed designs applied to FL. A few studies focus on FL's basic algorithms and propose methods to further increase their decentralization and scalability. In turn, a large part of this work aims to apply well-known distributed designs to the problem of FL. These include well-known peer-to-peer architectures as well as microservice-based fog architectures. Furthermore, these studies examine specific aspects of the FL architecture but focus on enabling their distribution and scalability.

These studies explore different architecture proposals to support the required communication and coordination features required by FL. A distributed environment based on Cloud and Fog computing concepts can provide the required scalability in an environment with a stable number of collaborating institutions. However, a fully distributed architecture, such as a BitTorrent swarm, enables a scalable solution to support dynamic scenarios like multiple institutions or individual patients collaborating on a global scale. Also, it can increase the reliability of the overall system by eliminating known problems such as single points of failure. However, a fully distributed architecture brings additional challenges that may require complex algorithms for management or decision making, for example. One particular challenge relevant to FL is the authenticity of transactions conducted throughout the training process. In that sense, blockchains represent a relevant architecture that natively provides authenticity guarantees to transactions. Even so, there are still guarantees necessary to FL with medical data requiring the implementation of specialized mechanisms. The following topics address work related to these specific aspects.

3.1.3 Aggregation Methodologies. Model aggregation is a fundamental step on FL. In summary, the aggregation step generates a global learning model by combining the parameters received from each participant in the FL process. The methodology used to select and combine parameters greatly influences the resulting global model accuracy. Consequently, several studies evaluate and propose solutions to improve the aggregation process and obtain models with higher accuracy.

Six studies aim to improve the aggregation results by including metrics that classify local models before being combined, enabling the algorithm to select only those that improve the resulting model. Boughorbel et al. [55] use the concept of uncertainty to classify the usefulness of local models. This property allows the aggregation algorithm to select only local models with high generalization to train the global model. Huang et al. [56] employ a cross-entropy analysis to evaluate the quality of local models before sending them to the aggregation server, which defines the threshold that characterizes a good model. The method reduces the computational and network requirements for training and improves the resulting model's accuracy. Cao et al. [57] propose a metric that quantifies the added knowledge from individual learning models in a federated context. The aggregation entity can then request only the local models with higher added knowledge to improve training accuracy and reduce network requirements. Chen et al. [58] evaluate the learning model parameters and propose a "sign" metric that indicates how much improvement an update contributes to the global model. Updates that decrease the global model accuracy are not submitted for aggregation. Similarly, Zhang et al. [59] propose a method to select which clients will participate in the global model training. Clients only upload the local model if it is guaranteed to contribute to the global model accuracy. Furthermore, the global model also discards updates received from clients that take too long to train a certain number of epochs for the model. Yang et al. [60] propose a method that leverages labeled and unlabeled data available in training datasets. Clients employ the current global FL model to generate pseudo-labels for unlabeled data, enabling it to improve the global model.

Other studies focus on the aggregation algorithm, evaluating and proposing alternatives to the traditional gradient descent, which is widely used in FL. Jiang et al. [61] propose the use of an attention mechanism to classify and select the parameters that will be aggregated in the global

model. Results indicate that the attention aggregation outperforms the gradient descent. Chen et al. [62] investigate the problem of activation divergence that occurs during the aggregation of models in FL, which can drastically increase their training time and reduce their accuracy. The authors propose a solution for this problem with a method that maximizes the entropy of activation vectors across all FL participants.

Two studies aim at modifying the local models to guarantee that dataset imbalances do not influence the resulting global model. Yan et al. [63] employ a **generative adversarial network (GAN)** to create a synthetic dataset that encompasses the representative characteristics of the original training data. The synthetic dataset eliminates data variations that can influence the global model while preserving the privacy of training data. Wu et al. [64] employ a generative convolutional autoencoder to guarantee class balance among distributed datasets used for in-home health monitoring.

Finally, two studies focus on improving the aggregation methodology to make it robust against malicious users trying to degrade the resulting global model. Su and Xu [65] propose a modification of the gradient descent method capable of tolerating input data from adversarial workers. The method can maintain learning model accuracy even with a fraction of malicious clients present in the infrastructure. In turn, Chen et al. [66] leverage the trusted execution environment available on modern computer platforms (e.g., Intel SGX) to verify and sign the results from individual contributions to FL. Such a method guarantees model authenticity during aggregation and prevents insider threats, patients' health data leakage, and model tempering at the aggregation engine and edge devices.

The articles in this category demonstrate that the main focus of work related to aggregation is to obtain the highest possible accuracy from the global model resulting from the process. To that end, three main lines of investigation exist: *(i)* applying metrics to rank and select only the models with higher knowledge derived from the training dataset; *(ii)* identify issues inherent from the distributed training process and propose solutions to reduce the deviation that it may cause; and *(iii)* mitigate possible attacks stemming from malicious participants trying to hinder the accuracy of the aggregated model.

The distributed nature of FL requires these additional methods to coordinate the training process since a centralized analysis of the entire dataset is not possible. Solutions that evaluate the accuracy and generality of local models before submitting them for global aggregation provide a solution without distributed coordination among peers. This characteristic reduces the network requirements from the overall infrastructure while reducing the impact of imbalanced local datasets. Nevertheless, local mechanisms still rely on local estimates that can deviate from the characteristics of the global dataset. In turn, solutions that exchange dataset quality information among FL peers can further improve the quality of local models before the aggregation process. However, these solutions impose additional network requirements to the architecture and may not be suitable for scenarios considering the participation of constrained devices (e.g., smartphones). Some proposals also consistently rely on using local datasets to verify the accuracy of local and global models. This approach may incur a verification bias similar to what may occur during model training with local datasets. Such analyses need to account for these biases, possibly using synthetic or sampled datasets, as some studies proposed. Overall, most studies show that FL imposes additional processing and communication costs to compensate for the distributed nature assumed by the training process.

3.1.4 Confidentiality Guarantees. Medical data is sensitive by nature. Consequently, the confidentiality of individual data items used to train models is one of the most stringent requirements for FL in medical applications. Usually, stewards use traditional mechanisms based on perimeter

security to ensure that only trustworthy parties have access to clinical data. However, scenarios with data sharing among potentially untrustworthy parties require different security mechanisms, such as the ones proposed by the zero-trust model [67]. The fundamental assumption of this model is that no network traffic is trustworthy until its proper authentication and authorization. It is compatible with modern system architectures, for example, distributed medical applications based on IoT [68]. This trust model is particularly relevant to FL applied to medical data because an institution or patient must distribute sensitive information to parties outside the boundaries of traditional trust enforcement mechanisms. As a result, no assumptions can be made about the data security of collaborating partners or communication channels. Furthermore, collaborating parties should assume that to avoid leakage of sensitive information, the more obfuscated it becomes during the FL process, the better.

Current studies on FL applied to medical data explore the assumptions of zero-trust models. Thus, they focus on evaluating and guaranteeing the confidentiality of training datasets to avoid leaking private data to untrustworthy parties. Early studies on the subject include the work from Hitaj et al. [69], which explores the use of a GAN to reconstruct the private data used to distributed train learning models. Their results show that a GAN can defeat most basic privacy mechanisms designed in traditional FL architectures and the need to develop stronger security mechanisms to protect medical data.

Three studies look into the application of individual security techniques to guarantee the confidentiality of model parameters and, as a result, the data used to train them. Bonawitz et al. [70] explore the use of secret sharing to secure model parameters before transmitting local models to collaborating parties. Dong et al. [71] further investigate the use of secret sharing and conduct an extensive evaluation on its confidentiality guarantees. Song et al. [72] analyze the efficacy of homomorphic encryption to protect learning models on FL applied to healthcare applications.

Differential privacy is a method that enables a given database to be modified without changing its properties and, thus, the results of statistical analyses conducted over the data [73]. It is a particularly promising technique to secure FL architectures. Li et al. [74] provide an in-depth analysis of differential privacy applied to FL for healthcare. Pfohl et al. [75] evaluate the accuracy of learning models trained with three systems: centralized, FL, and FL with differential privacy. Their results indicate that differential privacy negatively impacts the accuracy of learning models when their parameters are not correctly adjusted.

Furthermore, a considerable body of work explores the application of differential privacy combined with other techniques. Truex et al. [76] combine secure multiparty communication and differential privacy to guarantee the confidentiality of training data during the aggregation of local models. Zhang et al. [77] combine learning model partitioning and differential privacy to provide confidentiality guarantees on FL. The proposal aims to maintain a balanced trade-off between computational cost and privacy guarantees. Ma et al. [78] enhance tensor factorization [23] with differential privacy to guarantee the confidentiality of training data when applied to computational phenotyping of EHR. Triastcyn and Faltings [79] protect the parameters from local models with the Bayesian differential privacy technique. Results indicate that the method further reduces data leakage compared to traditional differential privacy, but it requires additional assumptions to be taken on the FL architecture. Li et al. [80] propose an FL architecture that combines differential privacy and the mixture of experts. The resulting architecture is resilient to inversion attacks on local models and degradation of model aggregation due to non-**independent and identically distributed (IID)** data. Gong et al. [81] propose to dynamically adapt the level of noise introduced with differential privacy methods applied to FL. The noise level varies with the significance of each neuron to the overall knowledge, which directly impacts the accuracy of the global model.

Some studies also focus on combining homomorphic encryption with other techniques to improve data and ML models' confidentiality. Such an approach allows data processing while both models and data remain encrypted on FL. Hao et al. [82] employ homomorphic encryption and augmented learning with error to protect the parameters of local learning models during training. The method reduces the network communication required to secure the aggregation process on FL without sacrificing model accuracy. Dong et al. [83] combine secret sharing and homomorphic encryption to protect the parameters from learning models shared in a federated architecture. The method also employs ternary gradients to encode model parameters to further reduce training data leakage.

Finally, a group of studies proposes techniques to manipulate learning models to reduce the chance of leakage during transmission. Cui and Liu [84] propose a hybridization method that recombines parameters from local learning models during aggregation and mitigates the possibility of the derivation of confidential data from the global learning model. Shao et al. [85] employ a stochastic channel-based method that identifies paths on the neural network with lower knowledge acquired during training. The method prunes these paths before sharing the model, reducing the chance of confidential data being derived. Ge et al. [86] propose an architecture that divides the learning model into private and public parts. The public part is trained with a subset of the private data to reduce data leakage during aggregation. In turn, the private part combines the global learning model shared by the central aggregation entity with additional training with the complete private dataset. Jeon and Kim [87] propose a method that splits local datasets into small batches individually sent to the aggregation entity. The method reduces the chance that the global model can be overfitted, which would enable the derivation of the training data from the model.

The literature analysis demonstrates that FL applied to EHR requires extensive work to mitigate risks to the confidentiality of data. To that end, studies explore various cryptographic techniques tailored for distributed applications. In particular, previous work demonstrates that homomorphic encryption [88] and differential privacy [89] can protect the datasets used for neural network training with minimal impact to model accuracy. These early results also indicate that these techniques are viable alternatives to mitigate possible attacks in federated learning architectures without compromising model accuracy. Results from studies combining different techniques also show that these techniques can complement each other to increase confidentiality guarantees in FL environments. The literature also shows that most studies focus on privacy guarantees on neural network models because they already offer an intrinsic level of privacy due to their non-interpretability. However, interpretable models can be a requirement for medical applications where the reasoning behind a given decision or diagnosis is a requirement. For such a scenario, a few studies presented promising results applying homomorphic encryption and differential privacy to train interpretable models such as random forests using FL.

3.1.5 Federated Data Analaysis. The final group of studies on the literature corpus focuses on the particular aspect of federated data analysis. These studies do not aim to generate a learning model using ML methods over distributed datasets. Rather, they aim to perform an initial analysis of the distributed data to facilitate the distributed learning procedure.

Some studies aim to process datasets with sensitive information and create anonymous versions of them to facilitate the distribution. Lee et al. [90] employ data hashing techniques to generate fingerprints of private databases and enable searches for specific parameters without compromising the confidentiality of the dataset. Their work enables institutions to explore a dataset prior to requesting private data for a data analysis study. Huang et al. [91] employ autoencoders and k-means clustering to generate an anonymous dataset that can be shared and grouped into a global structure that does not compromise confidential data. The resulting dataset enables distributed

model training with lower confidentiality requirements for communication mechanisms. Fioretto and Van Hentenryck [92] employ a two-step process to anonymize datasets that will be shared for ML. The process employs a predictor mechanism that describes all the parameters and data formats required by a specific learning task. This predictor is distributed to dataset owners and employed to generate anonymized datasets with the data already converted to the required format.

The second group of studies aims to enable the evaluation of distributed datasets to reduce data impurity prior to its usage for learning tasks. Liu et al. [93] employ a GAN to generate the missing data on local databases before use on an ML algorithm. The authors aim to improve the accuracy of local learning models used in the aggregation process of FL. Pezoulas et al. [94] propose an architecture for curation and federated analysis of medical datasets. Their goal is to enable dataset cleaning, model harmonization, and distributed analysis over data from multiple medical institutions without sacrificing confidentiality.

These studies demonstrate that there are issues beyond the scope of the FL process applied to medical and EHR data. The selected literature corpus presents work on generating anonymous copies of sensitive datasets to facilitate distribution and preparation prior to learning tasks. These studies aim to find solutions that enable FL peers to exchange information about local datasets and mitigate known problems such as class imbalances. They show promising results in achieving this goal without compromising data privacy. However, such methods may require additional data exchange among peers, thus increasing the overall network cost incurred by the FL process. Nevertheless, the possibility to align the private datasets to the specific requirements of learning tasks is essential to guarantee the quality of the locally trained models, which will directly impact the result of the overall process. Finally, results from these studies can be enhanced with privacy methods previously reviewed, such as homomorphic encryption.

3.2 Machine Learning Methods and Applications

Machine learning methods have a wide application in healthcare [95]. The results from our literature review also demonstrate this fact with the variety of case studies explored with FL. This section analyzes the machine learning methods used in studies about FL applied to EHR data and their main application. Table 2 summarizes the machine learning methods and applications observed in the literature corpus. Some articles did not describe a specific case study on the application of machine learning to healthcare data. Hence, the table excludes these particular studies.

Deep learning is the most common technique observed in the literature corpus, being used in almost 79% of the reported case studies. One particular area that benefited greatly from deep learning in computer vision. This trend is reflected in the literature corpus, in which all studies that explore medical image analysis employ deep learning. One area of particular interest is the analysis of neurological **magnetic resonance imaging (MRI)** data. Studies from Sheller et al. [24], Roy et al. [46], Li et al. [74], and Sheller et al. [33] employ FL to train models for segmentation of brain MRI data, in particular to identify regions with tumors. Some studies also propose to analyze MRI data to identify patterns representative of specific neurological conditions. In particular, Silva et al. [27] apply the ENIGMA Shape Analysis for prediction of neurological diseases in general, Lu et al. [45] employ shallow neural networks to diagnose Alzheimer's disease on patients with mild cognitive impairment, and Li et al. [80] use a multi-layer perceptron to identify autism spectrum disorder. Finally, Cao et al. [57] focus on medical images for cardiac diagnostics, proposing the use of deep neural networks to segment coronary vessels.

Another common application of deep learning with EHR data is the prediction of patient outcomes. In particular, studies from Liu et al. [25], Cui and Liu [84], Huang et al. [56], Pfohl et al. [75], Sharma et al. [29], Shao et al. [85], Gong et al. [81], and Vaid et al. [34] focus on predicting the mortality of patients using EHR data from ICUs. This case study is of particular interest to work

Table 2. Machine Learning Methods and Applications with EHR Data

	Deep Learning	Support Vector Machines	Bayesian Networks	Regression	Clustering	EHR Data Classification	EHR Data Extraction	Outcome Prediction	Disease Diagnostics	Medical Image Analysis
Kim et al. [23]	✓					✓				
Jochems et al. [22]			✓					✓		
Liu et al. [25]	✓							✓		
Shae and Tsai [51]	✓					✓				
Brisimi et al. [44]		✓						✓		
Sheller et al. [24]	✓									✓
Lee et al. [90]	✓					✓			✓	
Ma et al. [78]	✓					✓				
Lu et al. [45]	✓								✓	✓
Pfohl et al. [75]	✓							✓		
Boughorbel et al. [55]	✓							✓		
Liu et al. [26]		✓					✓		✓	
Huang et al. [91]	✓				✓	✓		✓		
Cui and Liu [84]	✓							✓		
Sharma et al. [29]	✓							✓		
Shao et al. [85]	✓							✓		
Liu et al. [93]	✓					✓			✓	
Roy et al. [46]	✓									✓
Sanyal et al. [47]				✓			✓			
Huang et al. [56]	✓							✓		
Silva et al. [27]	✓								✓	✓
Triastcyn and Faltings [79]	✓								✓	
Hao et al. [82]	✓						✓			
Lugan et al. [52]	✓						✓			
Truex et al. [76]	✓	✓						✓		
Fioretto and Van Hentenryck [92]		✓		✓					✓	
Jiang et al. [28]	✓						✓			
Li et al. [74]	✓									✓
Chen et al. [62]	✓					✓				
Ge et al. [86]	✓					✓	✓			
Ribero et al. [30]					✓			✓		
Cao et al. [57]	✓									✓
Li et al. [80]	✓								✓	✓
Pezoulas et al. [94]			✓					✓		
Deist et al. [31]				✓				✓		
Dong et al. [83]	✓						✓			
Gong et al. [81]	✓							✓		
Yan et al. [63]	✓									✓
Wu et al. [64]	✓							✓		
Hao et al. [50]	✓									✓
Chen et al. [58]	✓						✓			
Rahman et al. [54]	✓								✓	
Lim et al. [43]	✓					✓				
Guo et al. [38]	✓								✓	
Chen et al. [39]	✓								✓	
Ju et al. [40]	✓						✓			
Zhang et al. [41]	✓								✓	
Szatmari et al. [32]	✓					✓				
Sheller et al. [33]	✓								✓	✓
Zhang et al. [59]	✓								✓	
Yang et al. [60]	✓									✓
Xue et al. [48]	✓							✓		
Polap et al. [49]	✓								✓	
Can and Ersoy [42]	✓					✓				
Vaid et al. [34]	✓			✓				✓		
Lee et al. [35]	✓								✓	✓
Liu et al. [36]	✓							✓		

on FL because high-quality datasets are available to conduct experiments in a simulated federated environment. Such databases include the MIMIC-III [96] and the eICU [97], which are widely used for research on the application of ML for ICU EHR data analysis. Various of the above studies employ these databases with the same goal but present different experimental parameters (e.g., data division among nodes) related to the particular aspect of FL being investigated. One particular study, from Boughorbel et al. [55], employs deep learning to predict the outcomes of preterm birth using EHR data.

Some studies use deep learning to evaluate and classify the information contained in EHR data. Most of these studies aim to extract **ICD (international classification of diseases)** from EHRs and classify patients according to such data. For example, Shae and Tsai [51] and Chen et al. [62] employ deep learning to generate models to classify the records from an EHR database. The work from Kim et al. [23] and Ma et al. [78] focus on the computational phenotyping of EHR datasets to identify classifications of each health record. Finally, Ge et al. [86] employ a convolutional neural network to analyze unstructured medical texts and classify them according to multiple parameters (e.g., diagnosis and treatment). Some work goes further on the classification of EHR data and aims to achieve disease diagnostics. Lee et al. [90] and Liu et al. [93] employ EHR data from ICUs of multiple institutions to generate automated diagnostics models with deep learning. The work from Triastcyn and Faltings [79] focuses specifically on diseases related to blindness, employing a convolutional neural network to detect them.

One final group of studies that focuses on deep learning aims to use models for hand-writing [52, 82, 83] and speech [28] recognition. They do not focus specifically on analyzing EHR data, but they explore the usage of hand-written health records data to extract and interpret medical information.

The remainder of the literature corpus explores FL with other machine learning techniques. Jochems et al. [22] employ a Bayesian network to evaluate the post-treatment survival of cancer patients. Pezoulas et al. [94] use a Multinominal Naïve Bayes to determine the clinical outcome of lymphoma patients. Brisimi et al. [44] employ sparse support vector machines with binary supervised classification problems to predict hospitalizations due to cardiac events. Sanyal et al. [47] focus on time-series analysis techniques for real-time evaluation of data from IoT medical devices (e.g., multi-parametric monitors). Deist et al. [31] apply linear regression methods to predict the two-year survival of lung cancer patients. Finally, Ribero et al. [30] use K-means clustering to recommend treatments based on datasets about drug administration.

Regarding the combination of multiple techniques, the work from Liu et al. [26] employ a deep averaging network and SVM to predict diagnoses using plain text EHR records. Fioretto and Van Hentenryck [92] use linear regression and support vector machines for diabetes prediction. Finally, Truex et al. [76] explore multiple machine learning algorithms to classify EHR data from nurseries.

3.3 Discussion

This section builds upon the results observed so far in our literature analysis and draws some insights on the current state and future perspectives for FL applied to EHR data. Based on the data collected from our literature corpus, we propose a taxonomy to classify the observed work. Figure 3 illustrates the proposed taxonomy, including the percentage of studies that fall into each category.

The taxonomy reveals that deep learning is the most explored ML methodology when using FL with EHR data. As previously mentioned in our analysis, this is an expected result given the current popularity of deep learning methods and their variations. Furthermore, it is important to

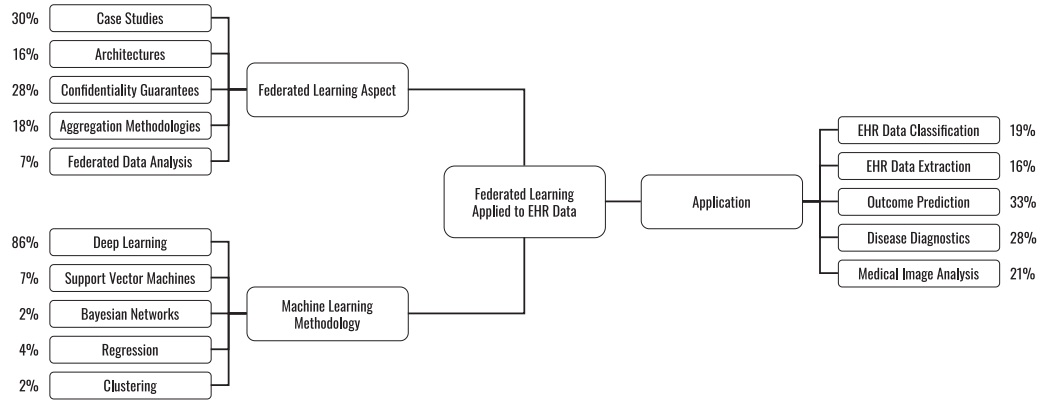

Fig. 3. Literature taxonomy for FL applied to EHR.

observe that various techniques employed in FL have high compatibility with deep learning-based models. For example, neural networks provide an intrinsic obfuscation to the data used for training because the interpretation of parameters is not a straightforward task. This fact makes methods based on neural networks ideal candidates for FL architectures. Nevertheless, as demonstrated by Hitaj et al. [69], this property does not automatically result in completely confidential models nor protects data while training and, thus, more advanced security mechanisms must be included in the architecture.

A few studies also explore other common learning models for medical applications, such as Bayesian networks or support vector machines. Some of these methods inherently expose the meaning of their parameters and require additional mechanisms to guarantee the confidentiality of sensitive medical data. On the other hand, these mechanisms enable the expansion of FL to a wide variety of ML techniques beyond deep learning, including other classical methods from medical data analysis. Such work requires an in-depth analysis of candidate ML methods and manipulating their parameters to obfuscate the respective training data.

Regarding the aspects from FL approached by different studies, there is a distinct interest in studies related to the confidentiality guarantees provided to the training data and resulting models. This interest is expected because of the high sensitivity of medical data and its regulatory protection requirements (e.g., GDPR or HIPPA). As a result, current work evaluates the degree of confidentiality provided by different FL approaches. Furthermore, it proposes mechanisms to mitigate parameter data leakage or model tampering that may expose protected information. These mechanisms focus on classical distributed confidentiality methods, such as differential privacy or secret sharing, to protect the distribution of learning models among collaborating partners. However, there are open questions on extending these cryptographic techniques to support ML methods such as decision trees, random forests, and KNN. These techniques may impose significant performance penalties for the ML models since they add another layer for security and privacy. Therefore, their performance needs to be improved dramatically.

Another subject of particular interest is the evaluation of case studies of EHR data analysis. This is a result of the novelty of the subject. There is a high interest in evaluating the advantages of FL applied to medical data and what scenarios present the most gains. These studies indicate that current case studies align with the traditional applications of ML to medical data, including prediction of patient outcomes, automated analysis of medical images, and automated diagnosis based on parameter analysis.

Current work presents an equal interest in architectures for FL, aggregation methods for model generation, and federated data analysis. The first two subjects are directly related to the communication cost and quality of the learning models extracted from the learning process. These subjects are of high relevance to FL architectures in general. Thus, it is expected that other subjects have a higher frequency in our results, given the focus of our literature search. Nevertheless, some studies approach these two subjects by looking at the specific characteristics of medical data and learning models. In turn, federated data analysis encompasses a series of studies that aim at evaluating and preparing datasets prior to the use of ML methodologies. This area is of particular interest because many applications require the analysis of distributed datasets to enable medical decision-making without the particular requirement of learning models. Consequently, there is space for further work on federated data analysis for medical data that can take advantage of the body of knowledge developed from work with FL.

Finally, regarding the applications of FL on EHR data, there is even interest in almost all subjects observed in the literature corpus. Outcome prediction, in particular, presented a higher interest than the average. As mentioned earlier, this results from the availability of high-quality public datasets that enable studies on outcome prediction based on ICU information. As a result, various works use such an application as a case study to evaluate FL mechanisms applied to medical data.

The observations discussed in this section stem from the analysis of the collected literature corpus. They enable us to better understand the current state of FL architectures applied in general and their requirements when EHR data is the focus of their application. Based on these insights, the following section discusses an overview of the main components of an architecture for FL for EHR data.

4 ARCHITECTURE FOR FEDERATED ANALYSIS AND LEARNING WITH EHR DATA

This section presents an architecture for federated analysis and learning with EHR data. The main goal of this architecture is to enable healthcare institutions (e.g., hospitals and laboratories) with access to private medical datasets to employ them in distributed data analysis and ML studies without compromising patient confidentiality. These institutions, defined as **data owners**, can maintain datasets with medical data of different characteristics, such as imaging exams, diagnostics results, and drug prescriptions. These datasets comprise sensitive information protected by laws and regulations, and individual patients should provide their explicit consent to make their data available for possible analysis in possible studies. We consider the acquisition of such consent the be an orthogonal issue to the design of the FL architecture, which can be tackled in future work. Nevertheless, we will consider mitigating threats to data privacy during data analysis or model training, as this is a key aspect of FL.

The architecture's design aims to leverage the best practices and methodologies observed in recent FL literature, as observed throughout Section 3. Figure 4 illustrates the components and their distribution among the different entities that take part in the learning process. These components are discussed in more detail next.

Data owners are presumed to be aware that participation in the FL infrastructure entails collaborating with privacy-preserving analyses and learning models. In exchange, the entity can request the execution of analysis and learning algorithms on other private entities to conduct medical data experiments. Other incentive mechanisms may exist within the FL framework, for example, financial or regulatory. Furthermore, current literature suggests that a blockchain infrastructure may be used as a collaborative infrastructure in which each partner becomes a node of the network and can use the learning models shared by other peers.

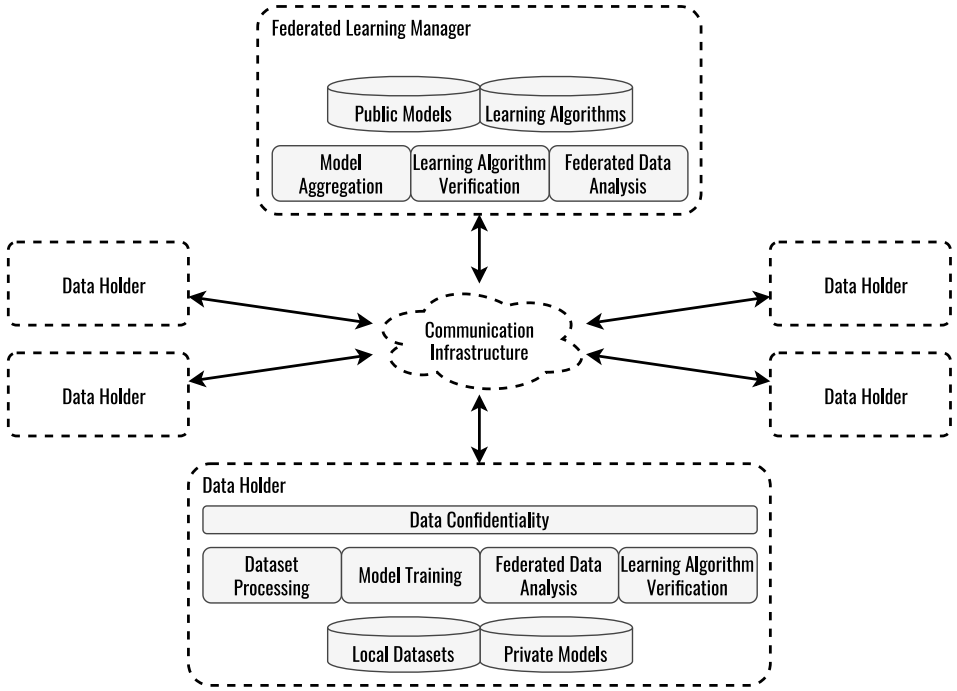

Fig. 4. Overview of an architecture for federated analysis and learning with EHRdata.

Due to the sensitive nature of the information contained in local datasets, the data owners will never directly share information with other parties. Furthermore, the data formats and standards used to store EHR data in each dataset can be disjoint among collaborators of the learning process. Each data owner uses their private datasets to train a **private model**, which can also include parameters obtained from models shared by other data owners from the FL manager. The data owner uses these private models to conduct internal data analyses and are not supposed to be directly shared with other FL entities.

Information sharing among data holders requires confidentiality guarantees for learning model parameters and analyses results. To that end, a **data confidentiality** module analyzes outgoing information (e.g., analysis results and learning models) to guarantee that it does not contain any sensitive information. To that end, this module can implement different methodologies based on a zero-trust security model [67]. Collaborating institutions will use public networks to exchange analysis results and learning models that may reveal private information. As such, security mechanisms must provide strong confidentiality guarantees, even if the institutions can establish trust among themselves for sharing medical data. Current literature shows that homomorphic encryption and differential privacy are techniques with substantial results to avoid data leaks without prohibitive computational costs [98]. If necessary, all collaborating parties should agree on the parameters of the confidentiality mechanisms (e.g., cryptographic keys).

The **dataset processing** module implements mechanisms that enable external learning algorithms to access and interpret the data contained in local datasets. Since the local data of each private entity may be in a particular format, tasks such as type conversion or image processing may be needed to enable a particular learning algorithm to process it. This module should

receive a template of the required data formats and identify the necessary conversions based on private data. To that end, a mechanism based on metadata can provide information on the properties of a dataset, enabling automated mechanisms to process data for compatibility with learning algorithms. The **learning algorithm verification** module implements mechanisms to analyze algorithms received from third parties and guarantee they will not purposely expose sensitive information from private datasets. One possibility to implement such verification is to include a "manifesto" with each learning algorithm that specifies input and outputs. Learning algorithms can then be checked to guarantee that they follow the manifesto specification. If a further guarantee is required, the manifesto can be human-verified to guarantee it does not violate any confidentiality restriction.

The **model training** module is responsible for executing the training algorithm, taking the data from the local datasets, and generating the resulting learning models. It executes both private and external learning algorithms, but, in the latter case, it presumes that the modules described above will verify the necessary conditions to avoid private data leakage. In turn, the **federated data analysis** module focuses on private or remote requests for data analyses that do not require learning algorithms. As mentioned before, federated data analysis is of particular interest to medical applications because it enables decision-making with broader knowledge about a given condition. The relevant security modules will also verify the external requests for federated data analysis to guarantee they will not leak any sensitive information. It is worth noting that data analysis or model training tasks require data holders to provide a certain amount of computational power to provide the required partial results. Thus, an incentive mechanism becomes necessary to justify this computational power. One straightforward incentive is the possibility of data holders using the analysis results to their own needs. However, such benefit requires that a given study provides relevant results to every data holder, which may be unfeasible. Thus, there is space for further development of incentive mechanisms for participation in FL studies, aiming at offsetting the costs incurred to data holders for knowledge generation.

All data holders communicate among themselves using a **communication infrastructure** that enables the secure exchange of learning models and analyses results through insecure networking environments. This is an important assumption given that the data holder will not be within the same administrative domain and must communicate using the Internet. The communication infrastructure also connects all data holders with the **federated learning manager**, which is the central element that coordinates the distributed learning and analyses processes. A common approach to implementing the manager is a centralized element (e.g., a cloud instance). However, decentralized approaches can also provide the functionality required to the architecture. For example, the literature corpus mentioned the use of swarm and blockchain infrastructures. Such peer-to-peer algorithms might be explored to deploy the complete functionality required from the public entity as part of the private entities themselves.

The manager implements the **Model Aggregation** methodologies required to collect all public models from private entities and merge them into a global learning model. Such methodologies will depend on the learning algorithms employed by the private entities. It is expected that the entity willing to train a certain learning model will provide the respective aggregation algorithm to the public entity. The federated data analysis module of the public entity focuses on the aggregation of data analyses results provided by private entities to generate the final analysis result. The public entity also implements its learning algorithm verification module, which can store and provide manifestos and participate in distributed analysis mechanisms requiring consensus among all participants. Current literature shows a variety of studies proposing methodologies to improve the aggregation process to improve the accuracy of global models. Results show that model

accuracy varies depending on the type of learning model and the characteristics of local datasets. As a result, support for different types of aggregation mechanisms may be necessary to guarantee compatibility with different learning algorithms.

The manager maintains a repository of **public models** generated with the aggregation methodology. Depending on the model's goals, it can be forwarded to the individual party that requested the training. They can also be redistributed to all participants in the learning process for their private model training or benchmarking with private data. The public entity also maintains a repository with all analysis and learning algorithms being processed by the infrastructure.

The description of the architecture and components presented above is not intended to be in-depth or complete. Our goal in this section is to provide an overview of the components that may be part of a more extensive infrastructure destined to analyze and learn EHR data in a federated fashion. Many elements required by these components are active research topics or still require studies to properly investigate them. As such, this description is intended to indicate which topics are still open issues in the viability of large-scale FL with EHR data.

5 FINAL REMARKS

Federated Learning presents itself as a promising solution to train machine learning models with large and diverse datasets without compromising information confidentiality. This characteristic is significant for healthcare applications, in which EHR data is very sensitive by nature and often cannot be easily shared. While there are some successful case studies for the use of FL in production scenarios, various research questions remain open before these architectures become available for widespread usage for healthcare applications.

This paper explored the landscape of research about FL applied to EHR data based on a literature corpus of 67 recent articles curated using a well-defined systematic review methodology. The analysis of these articles demonstrates various efforts to propose and improve methods to guarantee the confidentiality of training data. This result is straightforward given the stringent privacy restrictions related to the medical data from individual patients. Multiple studies also explore the application of FL on different types of medical data and how the resulting models compare against traditional training methods. Other topics include model aggregation methods tailored for the requirements of medical data and architectures to enable FL with multiple medical institutions. These insights indicate that FL applied to medical data is an active research area, and its fundamental aspects are still under development. It is important to note that the literature corpus analyzed in this study focuses on FL studies in healthcare contexts. However, FL techniques applied in other areas (e.g., Industry 4.0) may become relevant to the healthcare context in the future. Analyzing such articles and evaluating their applicability to healthcare scenarios is relevant. However, it lies beyond the scope of this review and is considered as future work.

Regarding research directions, the confidentiality of training data remains an open issue that requires further investigation and proposals. Particularly, there is space to explore confidentiality mechanisms for ML methods different from deep learning when used with federated training. There is space to investigate the effectiveness of aggregation methods with expanded characteristics, for example, with resilience to training problems on local models that may impact the accuracy of the global result. Finally, there is the need to further explore methods to enable the normalization of local datasets and to guarantee that different local datasets can provide the expected inputs and parameters to training models. Such dataset normalization used in an automated fashion is an open issue that is key for the widespread availability of FL in general.

REFERENCES

[1] D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, and G. Escobar. 2014. Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. *Health Affairs* 33 (2014), 1123–1131.

[2] J. Wiens and E. S. Shenoy. 2017. Machine learning for healthcare: On the verge of a major shift in healthcare epidemiology. *Clinical Infectious Diseases* 66 (2017), 149–153.

[3] N. H. Shah, A. Milstein, and S. C. Bagley. 2019. Making machine learning models clinically useful. *JAMA* 322 (2019), 1351–1352.

[4] S. Keyhani, P. L. Hebert, J. S. Ross, A. Federman, C. W. Zhu, and A. L. Siu. 2008. Electronic health record components and the quality of care. *JSTOR Medical Care* 46 (2008), 1267–1272.

[5] J. King, V. Patel, E. W. Jamoom, and M. F. Furukawa. 2014. Clinical benefits of electronic health record use: National findings. *Health Services Research* 49 (2014), 392–404.

[6] B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi. 2018. Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE Journal of Biomedical and Health Informatics* 22 (2018), 1589–1604.

[7] T. Chilimbi, Y. Suzue, J. Apacible, and K. Kalyanaraman. 2014. Project Adam: Building an efficient and scalable deep learning training system. In *11th USENIX Symposium on Operating Systems Design and Implementation (OSDI'14)*, 2014, pp. 571–582.

[8] E. Vayena, A. Blasimme, and I. G. Cohen. 2018. Machine learning in medicine: Addressing ethical challenges. *PLOS Medicine* 15 (2018), 1–4.

[9] E. Horvitz and D. Mulligan. 2015. Data, privacy, and the greater good. *Science* 349 (2015), 253–255.

[10] Q. Yang, Y. Liu, T. Chen, and Y. Tong. 2019. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technologies* 10 (2019), 12:1–12:19.

[11] R. Gu, C. Niu, F. Wu, G. Chen, C. Hu, C. Lyu, and Z. Wu. 2021. From server-based to client-based machine learning: A comprehensive survey. *ACM Computing Surveys* 54 (2021).

[12] A. Hard, K. Rao, R. Mathews, S. Ramaswamy, F. Beaufays, S. Augenstein, H. Eichner, C. Kiddon, and D. Ramage. 2018. Federated learning for mobile keyboard prediction. https://arxiv.org/abs/1811.03604. arXiv:1811.03604.

[13] P. Bellavista, L. Foschini, and A. Mora. 2021. Decentralised learning in federated deployment environments: A system-level survey. *ACM Computing Surveys* 54 (2021).

[14] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang. 2019. Federated learning for healthcare informatics. https://arxiv.org/abs/2003.08119. arXiv:1911.06270.

[15] N. Rieke, J. Hancox, W. Li, F. Milletari, H. Roth, S. Albarqouni, S. Bakas, M. N. Galtier, B. Landman, K. Maier-Hein, S. Ourselin, M. Sheller, R. M. Summers, A. Trask, D. Xu, M. Baust, and M. J. Cardoso. 2020. The future of digital health with federated learning. https://arxiv.org/abs/2003.08119. arXiv:2003.08119.

[16] F. Zerka, S. Barakat, S. Walsh, M. Bogowicz, R. T. H. Leijenaar, A. Jochems, B. Miraglio, D. Townend, and P. Lambin. 2020. Systematic review of privacy-preserving distributed machine learning from federated databases in health care. *JCO Clinical Cancer Informatics* 4 (2020), 184–200.

[17] L. Li, Y. Fan, M. Tse, and K.-Y. Lin. 2020. A review of applications in federated learning. *Computers & Industrial Engineering* 149 (2020), 106854.

[18] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava. 2021. A survey on security and privacy of federated learning. *Future Generation Computer Systems* 115 (2021), 619–640.

[19] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao. 2021. A survey on federated learning. *Knowledge-Based Systems* 216 (2021), 106775.

[20] H. Zhang and M. Ali Babar. 2013. Systematic reviews in software engineering: An empirical investigation. *Elsevier Information and Software Technology* 55 (2013), 1341–1354.

[21] T. Yang, G. Andrew, H. Eichner, H. Sun, W. Li, N. Kong, D. Ramage, and F. Beaufays. 2018. Applied federated learning: Improving Google keyboard query suggestions. https://arxiv.org/abs/1812.02903. arXiv:1812.02903.

[22] A. Jochems, T. M. Deist, I. El Naqa, M. Kessler, C. Mayo, J. Reeves, S. Jolly, M. Matuszak, R. Ten Haken, J. van Soest, C. Oberije, C. Faivre-Finn, G. Price, D. de Ruysscher, P. Lambin, and A. Dekker. 2017. Developing and validating a survival prediction model for NSCLC patients through distributed learning across 3 countries. *International Journal of Radiation Oncology, Biology, Physics* (2017), 344–352.

[23] Y. Kim, J. Sun, H. Yu, and X. Jiang. 2017. Federated tensor factorization for computational phenotyping. https://arxiv.org/abs/1704.03141. arXiv:1704.03141.

[24] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, and S. Bakas. 2019. Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. *Brainlesion* (2019), 92–104.

[25] D. Liu, T. Miller, R. Sayeed, and K. D. Mandl. 2018. FADL: Federated-autonomous deep learning for distributed electronic health record. https://arxiv.org/abs/1811.11400. arXiv:1811.11400.

[26] D. Liu, D. Dligach, and T. Miller. 2019. Two-stage federated phenotyping and patient representation learning. https://arxiv.org/abs/1908.05596. arXiv:1908.05596.
[27] S. Silva, B. A. Gutman, E. Romero, P. M. Thompson, A. Altmann, and M. Lorenzi. 2019. Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data. In *IEEE 16th International Symposium on Biomedical Imaging (ISBI'19)*. 270–274.
[28] D. Jiang, Y. Song, Y. Tong, X. Wu, W. Zhao, Q. Xu, and Q. Yang. 2019. Federated topic modeling. In *International Conference on Information and Knowledge Management (CIKM'19)*. 1071–1080.
[29] P. Sharma, F. E. Shamout, and D. A. Clifton. 2019. Preserving patient privacy while training a predictive model of in-hospital mortality. https://arxiv.org/abs/1912.00354. arXiv:1912.00354.
[30] M. Ribero, J. Henderson, S. Williamson, and H. Vikalo. 2020. Federating recommendations using differentially private prototypes 2020. https://arxiv.org/abs/2003.00602. arXiv:2003.00602.
[31] T. M. Deist, F. J. Dankers, P. Ojha, M. Scott Marshall, T. Janssen, C. Faivre-Finn, C. Masciocchi, V. Valentini, J. Wang, J. Chen, Z. Zhang, E. Spezi, M. Button, J. Jan Nuyttens, R. Vernhout, J. van Soest, A. Jochems, R. Monshouwer, J. Bussink, G. Price, P. Lambin, and A. Dekker. 2020. Distributed learning on 20 000+ lung cancer patients – the personal health train. *Elsevier Radiotherapy and Oncology* (2020), 189–200.
[32] T.-I. Szatmari, M. K. Petersen, M. J. Korzepa, and T. Giannetsos. 2020. Modelling audiological preferences using federated learning. In *28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP'20)*. 187–190.
[33] M. J. Sheller, B. Edwards, G. A. Reina, J. Martin, S. Pati, A. Kotrotsou, M. Milchenko, W. Xu, D. Marcus, R. R. Colen, and S. Bakas. 2020. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Nature Scientific Reports* 10 (2020), e24207.
[34] A. Vaid, S. K. Jaladanki, J. Xu, S. Teng, A. Kumar, S. Lee, S. Somani, I. Paranjpe, J. K. D. Freitas, T. Wanyan, K. W. Johnson, M. Bicak, E. Klang, Y. J. Kwon, A. Costa, S. Zhao, R. Miotto, A. W. Charney, E. Böttinger, Z. A. Fayad, G. N. Nadkarni, F. Wang, and B. S. Glicksberg. 2021. Federated learning of electronic health records to improve mortality prediction in hospitalized patients with Covid-19: Machine learning approach. *JMIR Medical Informatics* 9 (2021), e24207.
[35] H. Lee, Y. J. Chai, H. Joo, K. Lee, J. Y. Hwang, S.-M. Kim, K. Kim, I.-C. Nam, J. Y. Choi, H. W. Yu, M.-C. Lee, H. Masuoka, A. Miyauchi, K. E. Lee, S. Kim, and H.-J. Kong. 2021. Federated learning for thyroid ultrasound image analysis to protect personal information: Validation study in a real health care environment. *JMIR Medical Informatics* 9 (2021), e25869.
[36] J. C. Liu, J. Goetz, S. Sen, and A. Tewari. 2021. Learning from others without sacrificing privacy: Simulation comparing centralized and federated machine learning on mobile health data. *JMIR Mhealth and Uhealth* 9 (2021), e23728.
[37] C. A. da Costa, C. F. Pasluosta, B. Eskofier, D. B. da Silva, and R. da Rosa Righi. 2018. Internet of Health Things: Toward intelligent vital signs monitoring in hospital wards. *Artificial Intelligence in Medicine* 89 (2018), 61–69.
[38] Y. Guo, F. Liu, Z. Cai, L. Chen, and N. Xiao. 2020. FEEL: A federated edge learning system for efficient and privacy-preserving mobile healthcare. In *49th International Conference on Parallel Processing (ICPP'20)*. 1–11.
[39] Y. Chen, X. Qin, J. Wang, C. Yu, and W. Gao. 2020. Fedhealth: A federated transfer learning framework for wearable healthcare. *IEEE Intelligent Systems* 35 (2020), 83–93.
[40] C. Ju, D. Gao, R. Mane, B. Tan, Y. Liu, and C. Guan. 2020. Federated transfer learning for EEG signal classification. In *42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC'20)*. 3040–3045.
[41] M. Zhang, Y. Wang, and T. Luo. 2020. Federated learning for arrhythmia detection of non-IID ECG. In *IEEE 6th International Conference on Computer and Communications (ICCC'20)*. 1176–1180.
[42] Y. S. Can and C. Ersoy. 2021. Privacy-preserving federated deep learning for wearable IoT-based biomedical monitoring. *ACM Transactions on Internet Technology* 21 (2021).
[43] W. Y. B. Lim, S. Garg, Z. Xiong, D. Niyato, C. Leung, C. Miao, and M. Guizani. 2020. Dynamic contract design for federated learning in smart healthcare applications. *IEEE Internet of Things Journal* (2020), 1–10.
[44] T. S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. C. Paschalidis, and W. Shi. 2018. Federated learning of predictive models from federated electronic health records. *Elsevier International Journal of Medical Informatics* 112 (2018), 59–67.
[45] S. Lu, Y. Zhang, Y. Wang, and C. Mack. 2019. Learn electronic health records by fully decentralized federated learning. https://arxiv.org/abs/1912.01792. arXiv:1912.01792.
[46] A. G. Roy, S. Siddiqui, S. Pölsterl, N. Navab, and C. Wachinger. 2019. Braintorrent: A peer-to-peer environment for decentralized federated learning. https://arxiv.org/abs/1905.06731. arXiv:1905.06731.
[47] S. Sanyal, D. Wu, and B. Nour. 2019. A federated filtering framework for Internet of Medical Things. https://arxiv.org/abs/1905.01138. arXiv:1905.01138.
[48] Z. Xue, P. Zhou, Z. Xu, X. Wang, Y. Xie, X. Ding, and S. Wen. 2021. A resource-constrained and privacy-preserving edge computing enabled clinical decision system: A federated reinforcement learning approach. *IEEE Internet of Things Journal* (2021), 1–17.
[49] D. Polap, G. Srivastava, and K. Yu. 2021. Agent architecture of an intelligent medical system based on federated learning and blockchain technology. *Journal of Information Security and Applications* 58 (2021), 102748.

[50] M. Hao, H. Li, G. Xu, Z. Liu, and Z. Chen. 2020. Privacy-aware and resource-saving collaborative learning for healthcare in cloud computing. In *2020 IEEE International Conference on Communications (ICC'20)*. 1–6.
[51] Z. Shae and J. Tsai. 2018. Transform blockchain into distributed parallel computing architecture for precision medicine. In *IEEE 38th International Conference on Distributed Computing Systems (ICDCS'18)*. 1290–1299.
[52] S. Lugan, P. Desbordes, E. Brion, L. X. R. Tormo, A. Legay, and A. Macq. 2019. Secure architectures implementing trusted coalitions for blockchained distributed learning (TCLearn). *IEEE Access* 7 (2019), 181789–181799.
[53] F. Marulli, E. Bellini, and S. Marrone. 2020. A security-oriented architecture for federated learning in cloud environments. In *Web, Artificial Intelligence and Network Applications (WAINA'20)*. 730–741.
[54] M. A. Rahman, M. S. Hossain, M. S. Islam, N. A. Alrajeh, and G. Muhammad. 2020. Secure and provenance enhanced Internet of Health Things framework: A blockchain managed federated learning approach. *IEEE Access* 8 (2020), 205071–205087.
[55] S. Boughorbel, F. Jarray, N. Venugopal, S. Moosa, H. Elhadi, and M. Makhlouf. 2019. Federated uncertainty-aware learning for distributed hospital EHR data. https://arxiv.org/abs/1910.12191. arXiv:1910.12191.
[56] L. Huang, Y. Yin, Z. Fu, S. Zhang, H. Deng, and D. Liu. 2018. LoAdaBoost: Loss-based AdaBoost federated machine learning on medical data. https://arxiv.org/abs/1811.12629. arXiv:1811.12629.
[57] T.-D. Cao, T. Truong-Huu, H. Tran, and K. Tran. 2020. A federated learning framework for privacy-preserving and parallel training. https://arxiv.org/abs/2001.09782. arXiv:2001.09782.
[58] H. Chen, H. Li, G. Xu, Y. Zhang, and X. Luo. 2020. Achieving privacy-preserving federated learning with irrelevant updates over e-health applications. In *2020 IEEE International Conference on Communications (ICC'20)*. 1–6.
[59] W. Zhang, T. Zhou, Q. Lu, X. Wang, C. Zhu, H. Sun, Z. Wang, S. K. Lo, and F.-Y. Wang. 2021. Dynamic fusion-based federated learning for Covid-19 detection. *IEEE Internet of Things Journal* (2021), 1–8.
[60] D. Yang, Z. Xu, W. Li, A. Myronenko, H. R. Roth, S. Harmon, S. Xu, B. Turkbey, E. Turkbey, X. Wang, W. Zhu, G. Carrafiello, F. Patella, M. Cariati, H. Obinata, H. Mori, K. Tamura, P. An, B. J. Wood, and D. Xu. 2021. Federated semi-supervised learning for Covid region segmentation in chest CT using multi-national data from China, Italy, Japan. *Medical Image Analysis* 70 (2021), 101992.
[61] J. Jiang, S. Ji, and G. Long. 2020. Decentralized knowledge acquisition for mobile internet applications. *Springer World Wide Web* (2020).
[62] W. Chen, K. Bhardwaj, and R. Marculescu. 2020. FedMAX: Mitigating activation divergence for accurate and communication-efficient federated learning. https://arxiv.org/abs/2004.03657. arXiv:2004.03657.
[63] Z. Yan, J. Wicaksana, Z. Wang, X. Yang, and K.-T. Cheng. 2020. Variation-aware federated learning with multi-source decentralized medical image data. *IEEE Journal of Biomedical and Health Informatics* (2020), 1–14.
[64] Q. Wu, X. Chen, Z. Zhou, and J. Zhang. 2020. FedHome: Cloud-edge based personalized federated learning for in-home health monitoring. *IEEE Transactions on Mobile Computing* (2020), 1–14.
[65] L. Su and J. Xu. 2019. Securing distributed gradient descent in high dimensional statistical learning. *ACM SIGMETRICS Performance Evaluation Review* 47 (2019), 83–84.
[66] Y. Chen, F. Luo, T. Li, T. Xiang, Z. Liu, and J. Li. 2020. A training-integrity privacy-preserving federated learning scheme with trusted execution environment. *Elsevier Information Sciences* 522 (2020), 69–79.
[67] J. Kindervag, S. Balaouras, K. Mak, and J. Blackborow. 2016. *No More Chewy Centers: The Zero Trust Model of Information Security*, Technical Report, Forrester Research.
[68] B. Chen, S. Qiao, J. Zhao, D. Liu, X. Shi, M. Lyu, H. Chen, H. Lu, and Y. Zhai. 2021. A security awareness and protection system for 5G smart healthcare based on zero-trust architecture. *IEEE Internet of Things Journal* 8 (2021), 10248–10263.
[69] B. Hitaj, G. Ateniese, and F. Perez-Cruz. 2017. Deep models under the GAN: Information leakage from collaborative deep learning. In *SIGSAC Conference on Computer and Communications Security (SIGSAC'17)*. 603–618.
[70] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth. 2017. Practical secure aggregation for privacy-preserving machine learning. In *SIGSAC Conference on Computer and Communications Security (CCS'17)*. 1175–1191.
[71] Y. Dong, X. Chen, L. Shen, and D. Wang. 2019. Privacy-preserving distributed machine learning based on secret sharing. In *Information and Communications Security (ICICS'19)*. 684–702.
[72] L. Song, C. Ma, P. Wu, and Y. Zhang. 2019. PPD-DL: Privacy-preserving decentralized deep learning. In *Artificial Intelligence and Security (ICAIS'19)*. 273–282.
[73] C. Dwork. 2008. Differential privacy: A survey of results. *Lecture Notes in Computer Science* 4978 (2008), 1–19.
[74] W. Li, F. Milletarí, D. Xu, N. Rieke, J. Hancox, W. Zhu, M. Baust, Y. Cheng, S. Ourselin, M. J. Cardoso, and A. Feng. 2019. Privacy-preserving federated brain tumour segmentation. In *Machine Learning in Medical Imaging (MLMI'19)*. 133–141.
[75] S. R. Pfohl, A. M. Dai, and K. Heller. 2019. Federated and differentially private learning for electronic health records. https://arxiv.org/abs/1911.05861. arXiv:1911.05861.

[76] S. Truex, N. Baracaldo, A. Anwar, T. Steinke, H. Ludwig, R. Zhang, and Y. Zhou. 2019. A hybrid approach to privacy-preserving federated learning. In *12th ACM Workshop on Artificial Intelligence and Security (AISec'19)*. 1–11.
[77] J. Zhang, J. Wang, Y. Zhao, and B. Chen. 2019. An efficient federated learning scheme with differential privacy in mobile edge computing. In *Machine Learning and Intelligent Communications (MLICOM'19)*. 538–550.
[78] J. Ma, Q. Zhang, J. Lou, J. C. Ho, L. Xiong, and X. Jiang. 2019. Privacy-preserving tensor factorization for collaborative health data analysis. In *28th ACM International Conference on Information and Knowledge Management (CIKM'19)*. 1291–1300.
[79] A. Triastcyn and B. Faltings. 2019. Federated learning with Bayesian differential privacy. In *IEEE International Conference on Big Data (Big Data'19)*. 2587–2596.
[80] X. Li, Y. Gu, N. Dvornek, L. H. Staib, P. Ventola, and J. S. Duncan. 2020. Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: Abide results. *Medical Image Analysis* 65 (2020), 101765.
[81] M. Gong, K. Pan, Y. Xie, A. Qin, and Z. Tang. 2020. Preserving differential privacy in deep neural networks with relevance-based adaptive noise imposition. *Elsevier Neural Networks* 125 (2020), 131–141.
[82] M. Hao, H. Li, X. Luo, G. Xu, H. Yang, and S. Liu. 2019. Efficient and privacy-enhanced federated learning for industrial artificial intelligence. *IEEE Transactions on Industrial Informatics* (2019), 1–11.
[83] Y. Dong, X. Chen, L. Shen, and D. Wang. 2020. EaSTFLy: Efficient and secure ternary federated learning. *Elsevier Computers and Security* 94 (2020), 101824.
[84] J. Cui and D. Liu. 2019. Federated machine learning with anonymous random hybridization (FeARH) on medical records. https://arxiv.org/abs/2001.09751. arXiv:2001.09751.
[85] R. Shao, H. He, H. Liu, and D. Liu. 2019. Stochastic channel-based federated learning for medical data privacy preserving. https://arxiv.org/abs/1910.11160. arXiv:1910.11160.
[86] S. Ge, F. Wu, C. Wu, T. Qi, Y. Huang, and X. Xie. 2020. FedNER: Privacy-preserving medical named entity recognition with federated learning. https://arxiv.org/abs/2003.09288. arXiv:2003.09288.
[87] J. Jeon and J. Kim. 2020. Privacy-sensitive parallel split learning. In *International Conference on Information Networking (ICOIN'20)*. 7–9.
[88] N. Dowlin, R. Gilad-Bachrach, K. Laine, K. Lauter, M. Naehrig, and J. Wernsing. 2016. *CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy*. Technical Report, Microsoft Research.
[89] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang. 2016. Deep learning with differential privacy. In *2016 ACM SIGSAC Conference on Computer and Communications Security (CCS'16)*. ACM, 2016, 308–318.
[90] J. Lee, J. Sun, F. Wang, S. Wang, C.-H. Jun, and X. Jiang. 2018. Privacy-preserving patient similarity learning in a federated environment: Development and analysis. *JMIR Medical Informatics* 6 (2018), e20.
[91] L. Huang, A. L. Shea, H. Qian, A. Masurkar, H. Deng, and D. Liu. 2019. Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *Elsevier Journal of Biomedical Informatics* 99 (2019), 103291.
[92] F. Fioretto and P. Van Hentenryck. 2019. Privacy-preserving federated data sharing. In *18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS'19)*. 638–646.
[93] D. Liu, T. A. Miller, and K. D. Mandl. 2019. Confederated machine learning on horizontally and vertically separated medical data for large-scale health system intelligence. https://arxiv.org/abs/1910.02109. arXiv:1910.02109.
[94] V. C. Pezoulas, K. D. Kourou, F. Kalatzis, T. P. Exarchos, E. Zampeli, S. Gandolfo, A. Goules, C. Baldini, F. Skopouli, S. D. Vita, A. G. Tzioufas, and D. I. Fotiadis. 2020. Overcoming the barriers that obscure the interlinking and analysis of clinical data through harmonization and incremental learning. *IEEE Open Journal of Engineering in Medicine and Biology* 1 (2020), 83–90.
[95] B. Norgeot, B. S. Glicksberg, and A. J. Butte. 2019. A call for deep-learning healthcare. *Nature Medicine* 25 (2019), 14–15.
[96] A. E. Johnson, T. J. Pollard, L. Shen, H. L. Li-wei, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark. 2016. Mimic-III, a freely accessible critical care database. *Scientific Data* 3 (2016), 160035.
[97] T. J. Pollard, A. E. W. Johnson, J. D. Raffa, L. A. Celi, R. G. Mark, and O. Badawi. 2018. The eICU collaborative research database, a freely available multi-center database for critical care research. *Scientific Data* 5 (2018), 180178.
[98] B. Liu, M. Ding, S. Shaham, W. Rahayu, F. Farokhi, and Z. Lin. 2021. When machine learning meets privacy: A survey and outlook. *ACM Computing Surveys* 54 (2021).

Received March 2021; revised September 2021; accepted November 2021