



**HAL**  
open science

## Blockchain-Based Federated Learning in Medicine

Omar El Rifai, Maelle Biotteau, Xavier de Boissezon, Imen Megdiche, Franck Ravat, Olivier Teste

► **To cite this version:**

Omar El Rifai, Maelle Biotteau, Xavier de Boissezon, Imen Megdiche, Franck Ravat, et al.. Blockchain-Based Federated Learning in Medicine. International Conference on Artificial Intelligence in Medicine (AIME 2020), Aug 2020, Minneapolis, United States. pp.214-224, 10.1007/978-3-030-59137-3\_20 . hal-03080421

**HAL Id: hal-03080421**

**<https://hal.science/hal-03080421>**

Submitted on 18 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Blockchain-Based Federated Learning in Medicine

Omar El Rifai<sup>1,3</sup>(✉), Maelle Biotteau<sup>2,3</sup>, Xavier de Boissezon<sup>2,3</sup>,  
Imen Megdiche<sup>1</sup>, Franck Ravat<sup>1</sup>, and Olivier Teste<sup>1</sup>

<sup>1</sup> Institut de Recherche en Informatique de Toulouse SIG Team, Toulouse, France  
{omar.el-rifai, imen.megdiche, franck.ravat, olivier.teste}@irit.fr

<sup>2</sup> ToNIC, Toulouse NeuroImaging Center, Université de Toulouse, Inserm, UPS,  
Toulouse, France  
maelle.biotteau@inserm.fr

<sup>3</sup> Centre hospitalier universitaire de Toulouse, Toulouse, France  
deboissezon.xavier@chu-toulouse.fr

**Abstract.** Worldwide epidemic events have confirmed the need for medical data processing tools while bringing issues of data privacy, transparency and usage consent to the front. Federated Learning and the blockchain are two technologies that tackle these challenges and have been shown to be beneficial in medical contexts where data are often distributed and coming from different sources. In this paper we propose to integrate these two technologies for the first time in a medical setting. In particular, we propose a implementation of a coordinating server for a federated learning algorithm to share information for improved predictions while ensuring data transparency and usage consent. We illustrate the approach with a prediction decision support tool applied to a diabetes data-set. The particular challenges of the medical contexts are detailed and a prototype implementation is presented to validate the solution.

## 1 Introduction

Researchers face ethical challenges when handling medical records. Indeed, medical records hold sensitive information about patients that can be prejudicial if leaked. A recent controversy involving unconsenting access to tens of millions of identifiable health records re-sparked an interest in the data ethics debate [27]. As a consequence, medical institutions are reticent in sharing medical records [21]. Researchers go through time consuming procedures to request and use medical data-sets, often at the expense of efficiently advancing research. This situation is exacerbated for data scientists who use large, and heterogeneous data-sets scattered across different sites.

At the forefront of the ethical challenges we find, *data privacy*, *transparency*, and *usage consent* [15, 23, 27]. Data privacy is often thought in terms of identity privacy or confidentiality [23, 27]. Traditionally, anonymization techniques have been used for medical data processing [23]. Data transparency is about patients knowing and understanding how and by whom their data are used [27].

And usage consent means that data subjects have the right to decide how and when their data are used. [23,27]. Unfortunately, these points are often glossed over when processing patients’ data, and consent is sometimes implicitly assumed if data-sets are anonymized [15]. This situation is further complicated in the era of big-data where data-sets are scattered and processed by multiple often non-communicating parties [17].

Recently, two technologies emerged independently which address some of the issues highlighted above. Federated Learning (FL) emerged as a paradigm for training Machine Learning (ML) models across decentralized devices and minimizing the risk of exposing sensitive information [13,14,20]. While the blockchain emerged as a technology which offers unprecedented guarantees of reliability and usage transparency in decentralized settings [24].

This paper proposes a blockchain-based FL framework whereby the advantages of both technologies are put to use in the medical context. In particular, we propose a Smart Contract (SC) implementation of a coordinating server for a FL algorithm to ensure transparency and usage consent when sharing knowledge.

In Sect. 2, we discuss relevant related work for both FL and the blockchain in health care contexts. Then, in Sect. 3 we describe the problem in detail and explain how our solution can be applied for the medical setting. In Sect. 4, we show and discuss experiment results on a medical data-set for diabetes prediction. Finally, in Sect. 5, we present remaining challenges and open questions.

## 2 Health Care Analytics in Distributed Settings

Medical data offers a wealth of potential for improving the quality of care and reducing costs [21]. Nonetheless, medical data, as in other fields, is often “disorganized and distributed, coming from various sources and having different structures and forms” [21]. FL emerged as a response to these settings by providing a way to train ML models in heterogeneous and distributed settings while minimizing data transfers [13,14,20]. It has been used in health care settings because of the sensitive information of the data handled [32].

Likewise the blockchain has been studied in health care as it offers usage guarantees (transparency and immutability) not possible in traditional distributed data architectures [7,22]. In this section, we look at how these two technologies have been used in health care, and henceforth motivate the introduction of our blockchain-based FL medical decision support model.

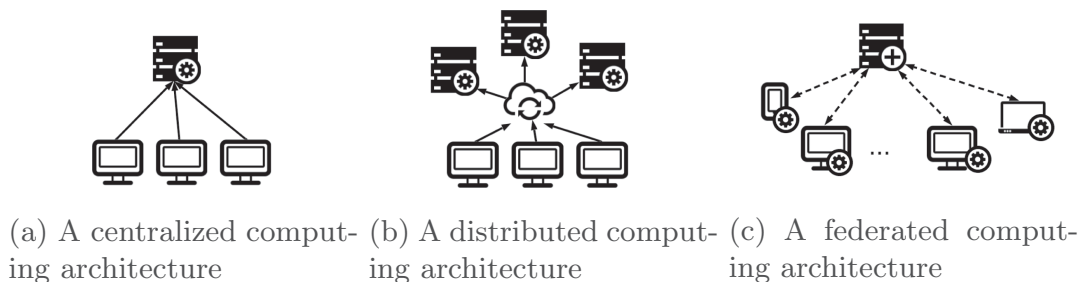
### 2.1 Federated Learning Approaches

Although works on distributed computing have been around for decades [18,19], new contexts have brought up previously unaddressed challenges: instead of evenly, and moderately distributed data-sets, FL approaches deal with uneven, and massively distributed data-sets [13,14,20]. These contexts have been shown to be applicable in health care settings to “connect all the medical institutions and makes them share their experiences with privacy guarantee” [32]. The term

FL was coined by Konevcny et al. [14] and McMahan et al. [20] which proposed a variant to the Stochastic Variance Reduced Gradient Descent (SVRGD) algorithm to solve ML problems in a distributed setting.

Characteristically, participants’ raw data never leaves the hosts’ devices in FL settings. The only data that is shared with a coordinating server are model parameters of local ML models. These parameters are then aggregated efficiently by a central server and the result sent back to the participants for updating their own models.

A comparison between the FL framework and more traditional computing architectures [13] is shown in Fig. 1. Figure (1a) shows an architecture where all computations are performed on a centralized server. In such an architecture, end devices query the server to use the computation model. The centralized server needs to hold the entire data-sets at the moment of the learning phase for this architecture to work and is typical of a siloed health facility. Figure (1b) shows a distributed computing architecture where multiple servers (sometimes defined as a cloud) share data-sets and the workload. This architecture is common for data-sets that have been anonymized and need to be shared across institutions. Finally, in Figure (1c), the end devices become active participants in the computation and only upload partial information to the server. The new role of the server is then to aggregate the information of the different devices and broadcast back the aggregated information to the end devices.



**Fig. 1.** An illustration of the different computing architectures. A gearwheel indicates where the main computation tasks is executed and dashed lines indicate that minimal information is transferred.

Studies that have used FL in the health care settings have been recently reported in [32]. For instance in [3], the authors develop a binary classification problem to predict cardiac-related hospitalizations based on data from electronic health records. They find that their algorithm converges faster than a centralized one at the expense of increased communication costs. In [11], the authors use a FL-based approach to predict mortality and hospital stay time. They improve on the baseline FL models by first clustering the patients into communities and outperform the baseline FL approaches.

However, in the words of the authors in [20]: “Clearly, some trust of the server coordinating the training is still required [when using FL]”. That is, server-side computations often suffer from a lack of transparency and it is difficult for users

to verify the computations performed [26]. Specifically, cloud computing has been argued to decrease the sovereignty of users over their data and models [5].

## 2.2 Blockchain with Health Care Applications

The blockchain was invented to solve the consensus problem in a decentralized, trust-less network [24]. That is, given a distrusting network of peers, the blockchain provides a mathematically robust way of verifying that data stored on our device are identical with data held by other peers. Initially designed with a financial application in mind, the blockchain quickly evolved to different domains [12].

In particular, the blockchain has been heavily applied in the medical domain whereby the properties it possesses (decentralization, immutability, and transparency) are core issues [1, 7, 22]. One such application which has been studied extensively is Electronic Health Records (EHR) management [2, 4, 16]. EHR management is inherently decentralized as stakeholders are distributed between patients, medical institutions and government institutions in some cases. Blockchain-based solutions, allow the different stakeholders to manager EHR transparently while guaranteeing fairness and usage (records access) consent [16].

## 2.3 Discussion

Both the blockchain and FL algorithms address important ethical challenges and have been successfully used in many health care settings [9, 32]. Indeed, by reducing the data that is shared when training models, FL algorithms reduce the risks of exposing sensitive patients data and hence address the privacy issue. Similarly, the blockchain addresses transparency and usage consent when dealing with medical records (c.f Appendix A.2).

Few papers have integrated ML models directly with the blockchain. Among those, Wang et al [29] for instance, have used the blockchain as a platform for hosting ML models, guaranteeing algorithmic correctness and usage traceability. Harris et al. [10] use the blockchain as a collaborative training platform for ML models. That is, the platform hosts ML models written in SCs and encourages data uploads from different users. Finally, the authors in [30] propose a domain agnostic setting where FL and blockchain technologies are successfully integrated. Indeed, a FL platform is used to train a deep learning network and instead of a centralized server, a SC is setup for federating the computations. The setup is said to provide “data confidentiality, computation auditability, and incentives for parties to participate in collaborative training” [30]. We bring these ideas for the first time to the medical context where privacy, transparency and usage consent are primordial. Indeed, medical records often fall under strict regulations such as the European General Data Protection Regulation (GDPR) and the suitability of such solutions needs to be investigated. Accordingly, we present in the next section, a medical decision support tool in a blockchain-based FL framework.

### 3 Problem Description

As discussed in previous sections, FL algorithms reduce the data communicated between participants and reduces the risks of exposing sensitive patients information compared to a cloud-based approach. Additionally, using SCs instead of a federating server, we can add transparency and usage guarantees to the setup. Specifically, at each federation round, the SC collects the values from the different participants and returns the aggregated parameters for participants to update their ML model in a completely automated way. Participants can also at any time verify the correct execution of this step in a completely transparent manner given the open and distributed execution of SCs. For this work, we choose to train an Artificial Neural Network (ANN) as previous results are available for comparison<sup>1</sup>. That said, other ML algorithms can easily be used in FL settings [14].

Let  $i$  be an index for the  $n$  different facilities that choose to participate in a collaborative ANN prediction model. The input to the model includes relevant patients' characteristics for the prediction model and the output is a binary variable. Each participant locally trains their ANN with weight parameters  $w$  and biases  $b$  [25] before sending the parameters to the federation SC. The loss function of such a ANN is of the following form.

$$f(w, b) \stackrel{def}{=} \frac{1}{2n} \sum_x \|y(x) - a\|^2 \quad (1)$$

Where  $x$  are the different data-set samples and  $y(x)$  corresponds to the output of our model for the particular input  $x$ . In order to train the model, the cost function [25] needs to be minimized so that the difference between  $y(x)$  and the actual output  $a$  (indicating if the patient of sample  $x$  is diabetic or not) is minimized. Of course, more complex objective function are possible but this is not the focus of this work.

The overall architecture of our system is illustrated in Fig. 2. At each round  $\tau$ , medical facilities send their trained parameters to a “Federating” SC uploaded on a blockchain network. The SC then aggregates the different parameters and sends back the result to the facilities for them to train the individual models again. As the aggregated parameters incorporate information from the different facilities, they have been shown to achieve near-optimal accuracy [20].

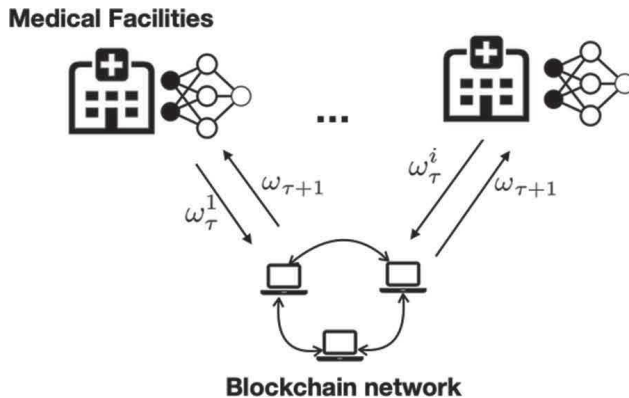
Let  $\omega$  represent both the weights and the biases, we use the formulation of Mc Mahan et al [20] in a federated setting and define the following loss:

$$\min_{\omega \in \mathbb{R}^d} f(\omega) \text{ where } f(\omega) \stackrel{def}{=} \frac{1}{n} \sum_{i=1}^n f_i(\omega) \quad (2)$$

The objective in Eq. 2 is to minimize the overall loss function defined as the average of the participants' individual loss functions. The algorithm for minimizing the loss functions can be separated into “client” steps and “server” steps

---

<sup>1</sup> <https://www.kaggle.com/ravichaubey1506/predictive-modelling-knn-ann-xgboost>.



**Fig. 2.** An illustration of the proposed system. The blockchain is used instead of the coordinating server in the original FL architecture.

as defined in [14, 20]. In our setting, the clients are the different facilities  $i$  and the “server” is the SC with access to all the local updates. The clients perform a normal gradient descent process and then send their parameters  $\omega$  to the SC which proceeds with a weighted average of the different client parameters as in Eq. 3.

$$\omega_{\tau+1} = \sum_i \frac{n_i}{n} \omega_{\tau} \quad (3)$$

Next, we investigate in Sect. 4 some experiments to validate the solution explained in this section to a real-world medical problem.

## 4 Experiments

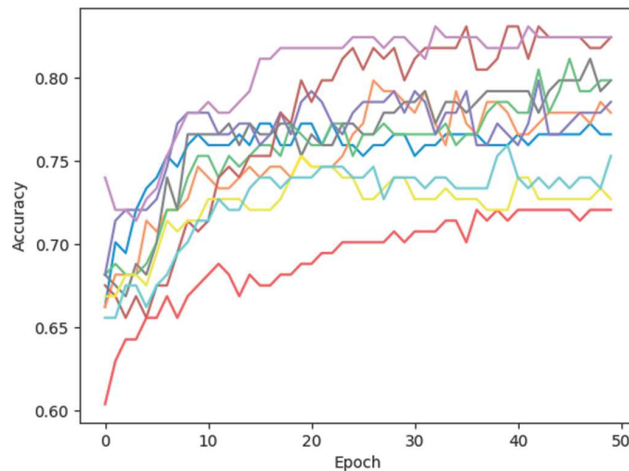
In this section we evaluate our solution using a diabetes data-set from the American National Institute of Diabetes and Digestive and Kidney Diseases available online<sup>2</sup>. First, we validate our ANN model’s capacity to predict diabetes by using it in a centralized setting. Then, we test the FL setting by distributing the medical records between 15 participants and having the participants collaborate. Finally, we implement a small blockchain prototype based on Ethereum to validate feasibility. For all the tests, we use ANN consisting of 2 hidden layers with 32 and 16 neurons respectively and a binary output layer to indicate whether the patients is predicted diabetic or not. Also, we use 80% of the data for training the ANN and the remaining 20% for testing.

In the first experiment the centralized model is trained for 50 epochs and the accuracy of the trained model on the testing data-set is evaluated after each epoch. Figure 3 illustrates the results obtained after running the experiment 10 times with different data distribution scenarios between the test and validation set. For each scenario, we randomly select 80% of the entries for training and leave the remaining for testing. We note that the initial data distribution

<sup>2</sup> <https://www.kaggle.com/uciml/pima-indians-diabetes-database>.



markedly impacts the accuracy of the final model. This clearly indicates an influence of the training data-set and corroborate the idea of benefiting from a collaborative setting.

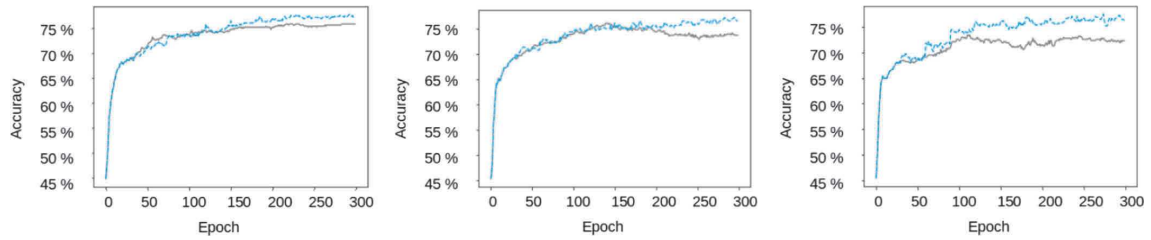


**Fig. 3.** The prediction accuracy of our ANN in a centralized setting for different training and testing scenarios.

Next, to test the FL model, we distribute the data-set among 15 participants to simulate a mid-sized collaboration scenario. After each 10 local epochs, the participants all share their model’s parameters (weights and biases) with a centralized federation server. The data is distributed among the participants in a randomized way. Some example results are shown in Fig. 4. The blue line represents the accuracy with the federated parameters while the grey line is a baseline model without federation. The results show that individual models’ accuracy are lower than in the centralized setting with an average accuracy of 73% (against 76% in the centralized setting) but all local models benefit from the data aggregates of the FL system with their accuracy improving between 1% to 5%. We choose to illustrate the results for 3 participants only as they had the most characteristic behavior, but other participants have similar trends. By having the ANN implemented in a FL setting, patients’ raw medical records are protected from attackers as they never leave the medical facilities. However, some studies have shown that inference attacks are possible on aggregated data. That is, some unintended information can be leaked from model parameters only. Possible mitigation strategies are discussed in Sect. 5.

Finally, with the above setup, participants have no visibility on the computations performed on the federating server. To add transparency and facilitate usage consent, we implement the federating algorithm in a SC using the Ethereum blockchain [31] and several open source tools (c.f Appendix A.4). Our implementation is openly accessible on this link <https://github.com/n-vcs/solidity-fl>. By having the SC deployed instead of a federating server, participants can verify how the parameters submitted are used and track all the transactions that are happening. Of course, a full-fledged implementation is needed to solve setting-specific issues that are not addressed in this paper due to space limitation.





(a) Low improvement with federation (b) Late improvement with federation (c) Significant improvement with the federation

**Fig. 4.** Resulting accuracy for three participants in a FL setting. (Color figure online)

**Clinical Relevance:** Promptly testing for diabetes can help patients receive timely and accurate treatment and help prevent misdiagnosis. Furthermore, training the model in the setting explained here allows medical centers to transparently and securely share knowledge with other facilities. This knowledge sharing creates more robust models which are resilient to environment-specific biases.

## 5 Conclusion and Future Work

In this work, we presented a blockchain-based FL solution to a diabetic prediction model. Our solution capitalizes on the benefits of both the blockchain and FL algorithms and caters to the particularities of a medical context. We provided a prototype implementation and presented tests to validate the relevance of the solution in the medical context. However, although no raw training data ever leaves the participants' devices in FL settings, the updates sent to the coordinating server may still contain private information. Indeed, FL algorithms have been shown to be vulnerable to inference attacks [30]. These types of attacks can deduce information about the training population which was not intended by the model.

To deal with this problem, privacy preserving protocols have been developed to prevent leakage of sensitive data. For instance, differential privacy solutions, Differential Privacy [6] and Homomorphic Encryption [8] have been shown to be relatively useful against inferential attacks [33]. However, even without revealing any details of the model, it is quite hard to achieve perfect privacy [28]. Against this grim outlook, it is however essential to be aware of the risks of each particular settings and measure them against the potential benefits.

**Acknowledgements.** This work was supported by a grant from the Roche Institute 2018.

## Appendix A The Blockchain Technology: Key Concepts and Implementation

This appendix explains some key concepts related to the blockchain as well as technical implementation details relevant to the present work.

### A.1 Main Characteristics

The blockchain is a distributed ledger technology managed by a network of peers. Data on the blockchain are visible and duplicated across participants. New data are added to the blockchain through a consensus protocol. As such, the blockchain is decentralized, immutable and transparent by design.

### A.2 Smart Contracts (SCs)

SCs are set of instructions, specified in digital form and executed when predefined conditions are met. Contrary to regular software, SCs benefit from the main characteristics of the blockchain and can help attain transparency and usage consent.

As opposed to centralized algorithms, SCs allow data owners to verify, at any time, the implementation of the FedAvg algorithm and replicate the results locally. Additionally, usage consent can be facilitated through SCs by logging data ownership certification directly on the blockchain as described in [16].

### A.3 Consensus and Incentive Mechanism

A consensus protocol is at the heart of the blockchain' mechanism. The Ethereum blockchain uses Proof-of-Work (PoW) by default. PoW relies on computational power to validate transactions or execute SCs. In our context, Medical institution are natural candidates for running the network as they will benefit from the resulting model. But to keep the network alive, nodes needs to be incentivized. As such, a possible set-up proposed in [30] is to have a reward mechanism for participants based on their data contribution.

### A.4 Current Set-up

For the prototype<sup>3</sup> an Ethereum blockchain is set-up on the back-end using the Truffle suite (Ganache and the Truffle Development Environment). Also, the FedAvg [20] algorithm and a basic ANN are developed using Python's scientific computing library (Numpy). The front-end is developed in React. Data for each participant are stored locally in a MongoDB database.

---

<sup>3</sup> <https://github.com/n-vcs/solidity-fl>.

## References

1. Agbo, C., Mahmoud, Q., Eklund, J.: Blockchain technology in healthcare: a systematic review. *Healthcare* **7**(2), 56 (2019)
2. Halamka, J.D., Andrew, M.D., Lippman A.E., Azaria, A.: A case study for blockchain in healthcare: “MedRec” prototype for electronic health records and medical research data. Technical report (2016)
3. Brisimi, T.S., Chen, R., Mela, T., Olshevsky, A., Paschalidis, I.C., Shi, W.: Federated learning of predictive models from federated electronic health records. *Int. J. Med. Inf.* **112**, 59–67 (2018)
4. Dagher, G.G., Mohler, J., Milojkovic, M., Babu, P.M.: Ancile: privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. *Sustain. Cities Soc.* **39**, 283–297 (2018)
5. De Filippi, S., McCarthy, S.: Cloud computing: Centralization and data sovereignty. *Euro. J. Law Technol.* **3**(2) (2012)
6. Dinur, I., Nissim, K.: Revealing information while preserving privacy. In: Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, pp. 202–210 (2003)
7. Drosatos, G., Kaldoudi, E.: Blockchain applications in the biomedical domain: a scoping review. *Comput. Struct. Biotechnol. J.* **17**, 229–240 (2019)
8. Gentry, C., Boneh, D.: A Fully Homomorphic Encryption Scheme, vol. 20. Stanford university Stanford, California (2009)
9. Gordon, W.J., Catalini, C.: Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability. *Comput. Struct. Biotechnol. J.* **16**, 224–230 (2018)
10. Harris, J.D, Waggoner, B.: Decentralized and collaborative AI on blockchain. In: 2019 IEEE International Conference on Blockchain (Blockchain), pp. 368–375 (2020)
11. Huang, L., Shea, A.L., Qian, H., Masurkar, A., Deng, H., Liu, D.: Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *J. Biomed. Inf.* **99**, 103291 (2019)
12. Janssen, M., Weerakkody, V., Ismagilova, E., Sivarajah, U., Irani, Z.: A framework for analysing blockchain technology adoption: integrating institutional, market and technical factors. *Int. J. Inf. Manag.* **50**, 302–309 (2020)
13. Kairouz, P., et al. Advances and open problems in federated learning. arXiv preprint [arXiv:1912.04977](https://arxiv.org/abs/1912.04977) (2019)
14. Konečný, J., McMahan, B., Ramage, D.: Federated optimization: distributed optimization beyond the datacenter. arXiv preprint [arXiv:1511.03575](https://arxiv.org/abs/1511.03575) (2015)
15. Larson, E.B.: Building trust in the power of “big data” research to serve the public good. *Jama* **309**(23), 2443–2444 (2013)
16. Leeming, G., Cunningham, J., Ainsworth, J.: A ledger of me: personalizing healthcare using blockchain technology. *Front. Med.* **6**(July), 1–10 (2019)
17. Leonelli, S.: Locating ethics in data science: responsibility and accountability in global and distributed knowledge production systems. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **374**(2083), 20160122 (2016)
18. Li, H., Ota, K., Dong, M.: Learning IoT in edge: deep learning for the internet of things with edge computing. *IEEE Netw.* **32**(1), 96–101 (2018)
19. Li, T., Sahu, A.K., Talwalkar, A., Smith, V.: Federated learning: Challenges, methods, and future directions. arXiv preprint [arXiv:1908.07873](https://arxiv.org/abs/1908.07873) (2019)

20. McMahan, B., Moore, E., Ramage, D., Hampson, S., et al.: Communication-efficient learning of deep networks from decentralized data. arXiv preprint [arXiv:1602.05629](https://arxiv.org/abs/1602.05629) (2016)
21. Mehta, N., Pandit, A.: Concurrence of big data analytics and healthcare: a systematic review. *Int. J. Med. Inf.* **114**, 57–65 (2018)
22. Mettler, M.: Blockchain technology in healthcare: the revolution starts here. In: 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services, Healthcom 2016, pp. 1–3 (2016)
23. Mittelstadt, B.D., Floridi, L.: The ethics of big data: current and foreseeable issues in biomedical contexts. *Sci. Eng. Ethics* **22**(2), 303–341 (2016)
24. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system. Technical report, Manubot (2019)
25. Nielsen, M.A.: *Neural Networks and Deep Learning*, vol. 2018. Determination press San Francisco, CA (2015)
26. Santhosh, G., De Vita, F., Bruneo, D., Longo, F., Puliafito, A.: Towards trustless prediction-as-a-service. In: Proceedings - 2019 IEEE International Conference on Smart Computing, SMARTCOMP 2019, pp. 317–322 (2019)
27. Schneble, C.O., Elger, B.S., Shaw, D.M.: Google’s project nightingale highlights the necessity of data science ethics review. *EMBO Mol. Med.* **12**(3), e12053 (2020)
28. Shokri, R., Stronati, M., Song, C., Shmatikov, V.: Membership inference attacks against machine learning models. In: 2017 IEEE Symposium on Security and Privacy (SP), pp. 3–18. IEEE (2017)
29. Wang, T.: A unified analytical framework for trustable machine learning and automation running with blockchain. In: Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018, pp. 4974–4983 (2018)
30. Weng, J., Weng, J., Zhang, J., Li, M., Zhang, Y., Luo, W.: DeepChain: auditable and privacy-preserving deep learning with blockchain-based incentive. *IEEE Trans. Dependable Secure Comput.* **8**, 1 (2018)
31. Wood, G., et al.: Ethereum: a secure decentralised generalised transaction ledger. *Ethereum Proj. Yellow Pap.* **151**(2014), 1–32 (2014)
32. Xu, J., Wang, F.: Federated learning for healthcare informatics. arXiv preprint [arXiv:1911.06270](https://arxiv.org/abs/1911.06270) (2019)
33. Yang, Q., Liu, Y., Chen, T., Tong, Y.: Federated machine learning: concept and applications. *ACM Trans. Intell. Syst. Technol.* **10**(2), 1–19 (2019)