Specifications for the Routine Implementation of Federated Learning in Hospitals Networks

Antoine Lamer, Alexandre Filiot, Yannick Bouillard, Paul Mangold, Paul Andrey, Jessica Schiro

Public Health and Informatics
J. Mantas et al. (Eds.)

doi:10.3233/SHTI210134

Specifications for the Routine Implementation of Federated Learning in Hospitals Networks

Antoine LAMER[a,b,1], Alexandre FILIOT[c], Yannick BOUILLARD[d], Paul MANGOLD[d], Paul ANDREY[c] and Jessica SCHIRO[a]
[a] *Univ. Lille, CHU Lille, ULR 2694 METRICS, F-59000, Lille, France*
[b] *Univ. Lille, Faculté Ingénierie et Management de la Santé, F-59000, Lille, France*
[c] *CHU Lille, INCLUDE, F-59000, Lille, France*
[d] *INRIA Lille Nord Europe, Magnet Team, F-59650, Villeneuve d'Ascq, France*

Abstract. We collected user needs to define a process for setting up Federated Learning in a network of hospitals. We identified seven steps: consortium definition, architecture implementation, clinical study definition, data collection, initialization, model training and results sharing. This process adapts certain steps from the classical centralized multicenter framework and brings new opportunities for interaction thanks to the architecture of the Federated Learning algorithms. It is open for completion to cover a variety of scenarios.

Keywords. Federated learning, Data privacy, Data reuse, Human factors

1. Introduction

In the usual analytical process involving multiple centers, patient data are shared to a centralized location to perform analyses, which raises privacy, technical, and data ownership concerns. As an alternative approach, Federated Learning (FL) consists in training statistical models in a decentralized way, leaving the data on each site, running computations locally and communicating aggregated information between centers during the training phase [1]. The two main topologies of networks for FL deployment come out of either an aggregation server (centralized) or a peer-to-peer (decentralized) workflow [2]. FL allows to use data from multiple centers, which often improves studies' robustness, while addressing the problem of data governance and improving data privacy.

In a previous study, we demonstrated that odd-ratios and their confidence intervals resulting from a decentralized logistic regression were consistent with the ones obtained with the traditional centralized model [3]. Other studies implemented Support-Vector Machine, Principal Components Analysis or Neural Networks using FL [4]. More recently, FL has been beneficial for predicting seven-days mortality in hospitalized COVID-19 patients from five local hospitals [5]. It may also be applied to patient similarity detection or adverse drug events detection [6]. Privacy-preserving

[1]Corresponding Author, Antoine Lamer, Univ. Lille, ULR 2694, 2 place de Verdun, F-59000, Lille, France; E-mail: antoine.lamer@univ-lille.fr

mechanisms, such as secure aggregation and differential privacy, can also be used to enhance FL so as to further address legal, ethical and security concerns related to data sharing [7].

Despite the significant spread of FL in the healthcare community, recommendations on how to deploy it in a routine use remain, to the best of our knowledge, scarce. In this paper, we propose an analysis of the existing situation to submit specifications to efficiently design and implement a federated learning framework across a network of hospital partners, based on the needs of hospital actors themselves.

2. Methods

After identifying all the actors involved in multicenter clinical studies, semi-structured interviews were performed by phone call in 7 French university hospitals (Amiens, Caen, Dijon, Lille, Marseille, Rouen, Toulouse). Interviewees were researchers, physicians, IT engineers, data scientists and data managers. The interviews aimed to understand how to implement FL in routine, to identify which actors are involved at each stage, their function and points of concern, along with their ideas and proposals to improve this process practice. Specific questions were then asked according to the profiles of the interviewees (Table 1).

Table 1. Interview questions

Actors	Description
All	How and by who are multicenter studies implemented today? What difficulties are you typically encountering? How do you imagine using federating learning in your multicenter studies? In your opinion, what are the technical and human resources required for FL? Do they differ from those required for multicenter studies?
Clinical researchers	How are the centers invited to join a multicenter study? How are research questions submitted to potential partners? Which kinds of information do partners share? What are the main steps for validating a multicenter protocol? Which legal constraints do you encounter during the process? Which results are returned to the partners in classical multicenter studies? How would you like to receive and/or visualize the results?
Data managers	How and by who is the data sent to the investigating center? Should it be the same person in a federated setting? Is the data stored in a central repository? In which format?
Data scientists	How to set the learning algorithms' hyperparameters? How to verify data integrity and integrity of computation?
IT engineers	What kind of technical architecture can be provided? Where can data be stored to be accessible? Does such a technical system raise any issues for getting legal authorities' approval? How are security audits usually performed on the system?

3. Results

We interviewed 18 people from 7 French university hospitals: 3 from Amiens, 1 from Caen, 1 from Dijon, 4 from Lille, 3 from Marseille, 5 from Rouen and 1 from Toulouse. We identified seven stages for deploying a FL workflow: consortium definition,

architecture implementation, clinical study definition, data collection, initialization, model training and results sharing. Figure 1 describes this sequence along with the actors involved at each stage. Thereafter, it will be admitted that the initiative for multicenter studies using FL is carried out by the "investigating center". By definition, this requires the investigators to have experience related to FL techniques.

3.1 Setting up the consortium

The following two steps need to be performed only once:

Consortium definition (decision-makers, researchers): search for collaborators (*e.g.* some research teams in public hospitals working on similar fields, or members of any biostatistical department) that might be part of the consortium. A short introduction with the basics of federated learning, current applications and added-value should be given to centers with no prior knowledge on FL. In particular, one should give incentives to centers that might be reluctant to go into decentralized multicenter studies using FL. As part of prospecting for collaborators, one should identify the different actors that might be involved in each center (*e.g.* IT engineers, researchers, physicians, legal experts, data protection officers), with a point of vigilance for the IT engineer who is the key person in the process. Once centers (and corresponding actors) have been identified, they should agree on a collaboration policy and address all legal questions.

Architecture implementation (IT engineers, data scientists): deployment of the logistic framework that will allow the different centers to communicate and the future studies to be launched easily without resorting to the IT engineers (a data scientist could possibly deal with IT issues instead). This step notably implies setting up virtual machines with secured communication ports and running tests to ensure the readiness of the FL workflow.

3.2 Performing the study

The following five steps are to be run for each study:

Clinical study definition (researchers, physicians, data scientists):

(a) Proposition of a study by some center(s) to the consortium with a principal and potentially secondary objectives. According to the pursued objectives, an overview of the data needed should be provided as a first simple eligibility criterion : in center c, does any research team, hospital department, individual researchers or doctors have the desired data and would be interested by the aforementioned objectives? does the center's data warehouse hold the targeted data? This way, collaborators can be identified in each center.

(b) Establishment of a detailed research protocol (collectively or by some centers). As in multicenter studies, objectives, data, methods and results dissemination are to be agreed upon beforehand. However, in federated computations, methodological details such as statistical model, learning algorithms and data integrity tests must be anticipated. Technical and practical organization points should also be discussed: how to guarantee privacy? when to execute the computations? how to handle unstable networks? can new partners be added during computations?

Data Collection (data scientists, researchers, physicians): prospective or retrospective center-wise data collection. An interoperable data set is made accessible to the FL algorithms, provided it can lawfully be used for research purposes.

Initialization (IT engineers, data scientists): data integrity and secure network connections verification. To check for outliers, verify model assumptions and/or to normalize data sets, a few statistical estimators may be exchanged across centers.

Model training (IT engineers, data scientists): when all participants are ready, the model training is launched by the investigator. Local gradients are computed on each center's data, and then exchanged across the network in order to update the shared models; this is repeated until model convergence, according to the criteria agreed upon in the protocol. At the end of this step, a validation check should be run to assert that all centers share the same results.

Results sharing (researchers, physicians): each participant retrieves the results of the study. Those can be measurements (odd-ratios, confidence intervals, F1 score, p-values, ...), models' parameters or models' predictions. Models' losses as well as system logs collected during the previous step should be analyzed by data scientists and IT engineers respectively. Any system dysfunction that could compromise the results should be reported to the investigator. Optionally, participants may be asked to validate results to allow moving on to further analyses and dissemination.

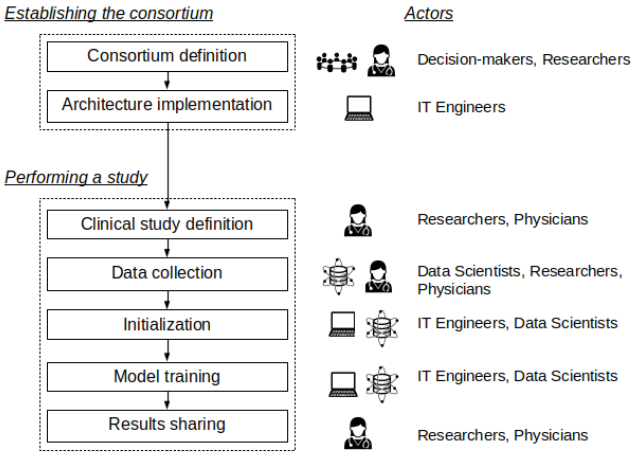

Figure 1. Proposed steps of a federated learning process in a consortium of hospitals.

4. Discussion

In this paper, we interviewed public hospitals' actors from 7 French centers. The objective was to analyse user needs for the design and the effective deployment of federated multicenter clinical studies. As part of a generic roadmap, we identified seven key steps designed for any member of the healthcare community willing to go into FL. One advantage of the proposed process is to keep close to the well-established process of centralized multicenter studies. Indeed, the FL process shares most of the classical steps in place in clinical research centers: consortium setup, study proposal and acceptance, results dissemination.

We remark that in contrast with classical multicenter studies, which only require partners to collect data and may fully delegate their processing to the investigator in charge of their statistical exploitation, FL relies on centers' ability to handle some data

preparation locally, and to allocate computational resources for model training, which can engage supplementary costs. In FL, centers also need to be available simultaneously at both computation and evaluation times, with reliable communication channels. However, distributing the study definition, modeling and evaluation steps yields increased and more diverse opportunities for data-owners to review and validate the methodology and results. This empowers non-investigating centers and can potentially result in better-controlled study quality.

Our framework remains general, leaving several points (*e.g.* network topologies, types of study, classes of statistical models...) open for refinement into a variety of scenarios. This work is therefore a first step towards a full specification of a federated learning protocol and its implementation in different settings. In our opinion and personal experience, those guidelines may be useful to anyone interested in FL from a statistical and/or clinical point of view, and willing to design new decentralized studies from scratch. We did not address the choice of a specific tool or library to perform federated computations. We note that existing tools [3,8] could fit within our framework, but usually only tackle the model training step, and are for the most still limited to simulated network environments.

5. Conclusions

In this study, we propose specifications for setting up and routinely running FL in a network of hospitals. It adapts certain steps of the multicenter centralized process and brings new opportunities for interaction, data ownership and flexibility thanks to the FL architecture. This general process is open to completion, so as to cover with greater precision the plurality of scenarios it can be applied to.

References

[1] Sheller MJ, Edwards B, Reina GA, Martin J, Pati S, Kotrotsou A, Milchenko M, Xu W, Marcus D, Colen RR, Bakas S. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. Sci Rep. 2020 Jul 28;10(1):12598.

[2] Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. NPJ Digit Med. 2020;3:119.

[3] Mangold P, Filiot A, Moussa M, Sobanski V, Ficheur G, Andrey P, et al. A Decentralized Framework for Biostatistics and Privacy Concerns. Stud Health Technol Inform. 23 nov 2020;275:137-41.

[4] Xu J, Glicksberg BS, Su C, Walker P, Bian J, Wang F. Federated Learning for Healthcare Informatics. J Healthc Inform Res. 12 nov 2020;1-19.

[5] Vaid A, Jaladanki SK, Xu J, Teng S, Kumar A, Lee S, Somani S, Paranjpe I, et al. Federated Learning of Electronic Health Records Improves Mortality Prediction in Patients Hospitalized with COVID-19. JMIR Med Inform. 2020 Dec 14.

[6] Choudhury O, Park Y, Salonidis T, Gkoulalas-Divanis A, Sylla I, Das AK. Predicting Adverse Drug Reactions on Distributed Health Data using Federated Learning. *AMIA Annu Symp Proc*. 2020;2019:313-322. Published 2020 Mar 4.

[7] Kaissis GA, Makowski MR, Rückert D, *et al.* Secure, privacy-preserving and federated machine learning in medical imaging. Nat Mach Intell 2; 305-311. Published 2020 Jun 8.

[8] Ryffel T, Trask A, Dahl M, Wagner B, Mancuso J, Rueckert D, Passerat-Palmbach J. A generic framework for privacy preserving deep learning. PPML 2018. arXiv: 1811.04017