DEPLOYING FEDERATED LEARNING ACROSS FRENCH HOSPITALS – LESSONS LEARNED

Aurélien Bellet (Inria)

A collaboration between Inria Magnet, CHU Lille (Include team) and GCS G4

Workshop “Data et Services pour une Ville Durable” (DSVD)
January 13, 2023

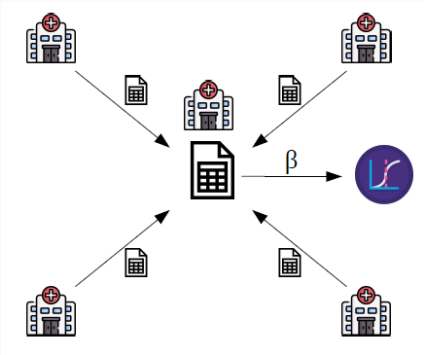
The Flamed Project

FLAMED project (started in 2020):
deploy Federated Learning (FL) approaches
to run multi-centric medical studies across 4 French hospitals

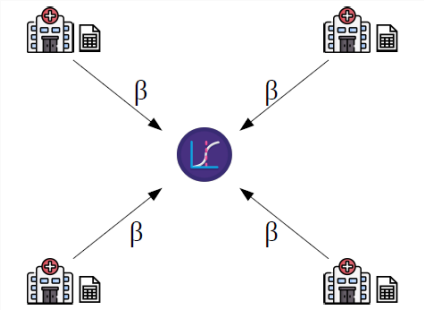
- A natural application for the fundamental work we do in Inria Magnet on federated, decentralized and privacy-preserving machine learning
- A response to strong incentives for hospitals to exploit retrospective data (creation of local data warehouses) while addressing restrictions on health data sharing
- An alternative to national centralized initiatives (e.g., Health Data Hub) in which hospitals lose control of their data

CENTRALIZED VS FEDERATED MULTI-CENTRIC STUDY

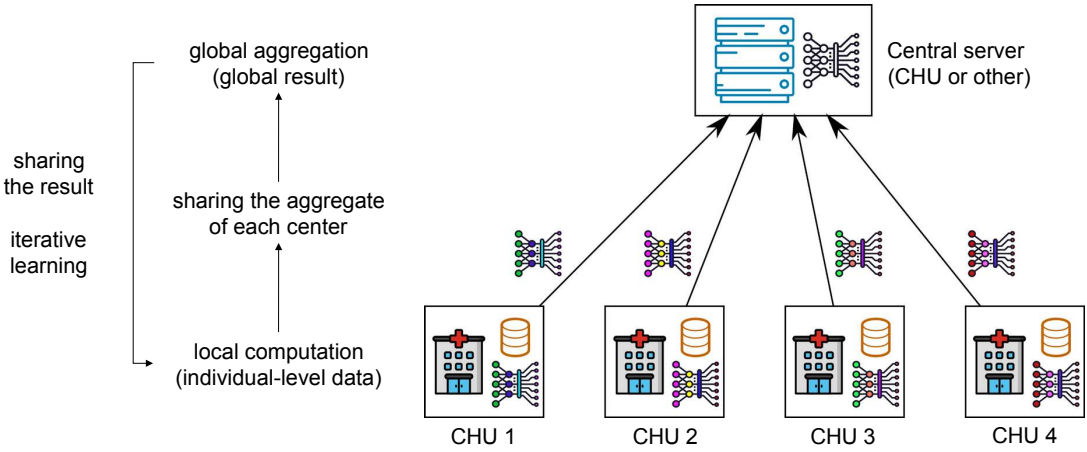
Centralized multi-centric study



Federated multi-centric study



FEDERATED MULTI-CENTRIC STUDY: GENERAL PRINCIPLE



WHAT WE HAVE ACHIEVED SO FAR

- Development of `DecLearn`, our own open-source Python library (ongoing work)
  - Choice of an in-house solution for maximum flexibility
  - Generic model API to easily accommodate any ML backend (e.g., scikit-learn, Pytorch...)
  - Modular algorithm API to easily implement advanced FL algorithms
  - Implementation of local differential privacy
  - Server-client communication with gRPC
  - Short term plan: releasing code and integration of some features into Fed-BioMed
- Proof-of-concept deployments across several hospitals
  - Current scope: the 4 CHUs of GCS G4 (Lille, Rouen, Caen, Amiens)
  - Synthetic and public data only
- Study of the GDPR requirements to deploy on real hospital data
  - Project selected for the digital health sandbox of CNIL (France's Data Protection Authority)
  - 2 real use-cases: prediction of re-admission, and diagnostic coding from medical reports
  - Aspects related to legal classification of aggregate quantities, obtaining consent, DPIAs...

Some Lessons Learned

ON THE INTERACTION WITH MANY STAKEHOLDERS

- During the course of the project we have interacted with:
  - Clinicians: identify relevant and feasible use-cases, understand their usual workflow, find the right arguments to sell FL (not always the ones you expect initially!)
  - DPOs: validate protocols internally, help with applications to external regulation bodies
  - IT people: get virtual machines, understand (typically heterogeneous) security policies
  - Engineers: export data from warehouse, deploy the FL solution and retrieve results/logs
  - Management people: create institutional/political incentives to make things happen (it helps a lot if the involved institutions have preexisting incentives to collaborate)
- In our case we benefited **a lot** from the fact that CHU Lille has a dedicated data warehouse team with people that understand both technical machine learning and biomedical aspects, know who to ask for something, etc

ON PRIVACY REQUIREMENTS: LEGAL VS TECHNICAL

- In GDPR (and other legal text), the notion of personal data is vague and subject to interpretation: privacy requirements are enforced through risk (self-)assessment and liability → can be unsettling for a scientist!
- Evaluating privacy risks of machine learning is notoriously difficult, but can be possible in a "best effort" sense (which is what GDPR requires)
- There is a need to come up with a risk assessment methodology for FL that can apply (or be easily adapted) to various use-cases, e.g., building on recent privacy attacks
- The use of actual PETs (e.g., differential privacy, secure aggregation, homomorphic encryption) is not always necessary (simpler measures can sometimes sufficient) but in critical cases they can help mitigate risks to an acceptable level

ACKNOWLEDGMENTS

Many thanks to:

- **Current and past members of Magnet:** Paul Andrey, Nathan Bigaud, Yannick Bouillard, Paul Mangold, Marc-André Sergiel, Marc Tommasi, Nathalie Vauquier
- **Our collaborators** at CHU Lille, Rouen, Caen and Amiens