

Fed-BioMed

Deploying Federated Learning in Real-World Health Applications

Marco Lorenzi

Inria Sophia Antipolis, Université Côte d'Azur

Epione Research Group

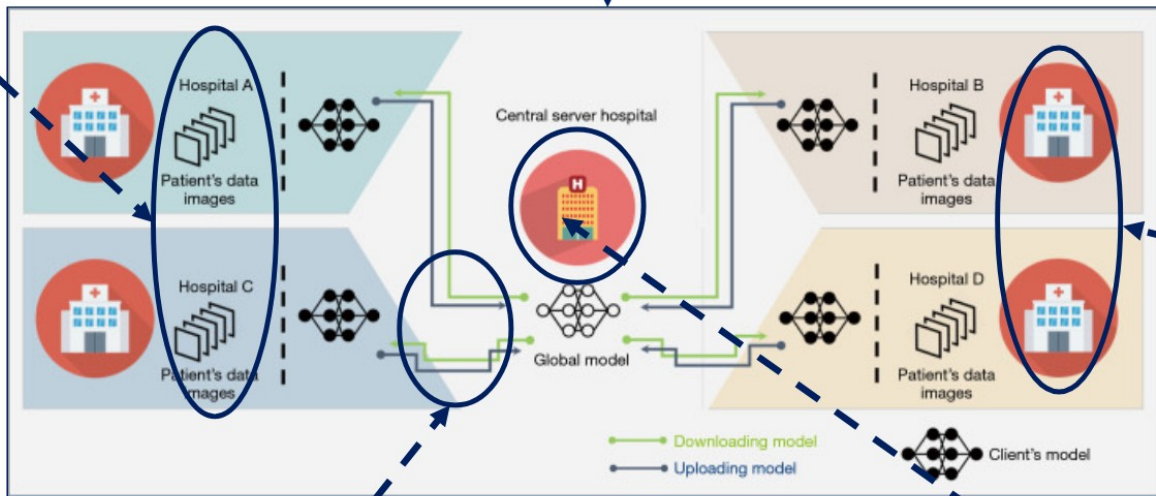


The Challenge of Medical Applications

Heterogeneous and specialized networks / database / data format

Security model

Regulation:
Data Protection
Officer,
Homologation,
Ethical Committee

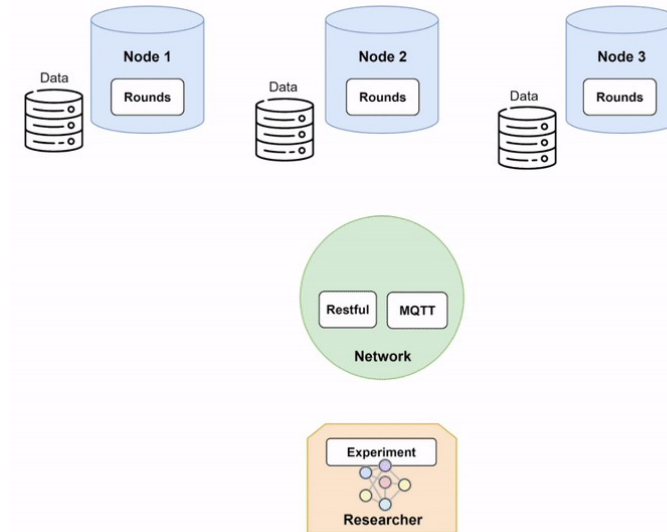


Availability of local
resources
(hardware and
personnel)

Data flow

Governance:
Server location,
Researcher role,
Hospital control over FL

Fed-BioMed: Federated Learning for Healthcare



Website: <https://fedbiomed.gitlabpages.inria.fr/>
GitLab: <https://gitlab.inria.fr/fedbiomed/fedbiomed>

Federated learning principles:

- data privacy;
- distributed training and testing.

Governance and data control:

- clinical data providers control their data sharing patterns;
- simplify usage of clinical-user-facing tools;
- review, audit and veto local training.

Interoperability of medical data:

- easy-to-use interfaces with widely-used data standards.

Interactivity:

- data scientists have fine-grained control over the execution of a federated training experiment;
- real-time monitoring and experiment steering between FL rounds, checkpointing.

Security and cybersecurity:

- protect data and models from cyberattacks and model inversion;
- secure, encrypted infrastructure;
- differential privacy.

Empowering scientific research:

- interactive, exploratory workflows, not industry-grade deployment;
 - collaborative research efforts, encourage exchange and feedback from users.
-

Programming languages in Fed-Biomed

- Python (AI frameworks: Sklearn, PyTorch, MONAI; GUI backend: Flask; jupyter-notebook)
- Javascript / react (for GUI)
- bash/script
- C++ (MP-SPDZ for secure aggregation)
- Docker (or singularity)
- Network components (MQTT, gunicorn, django for data exchange)
- Wireguard (VPN)

FL Security

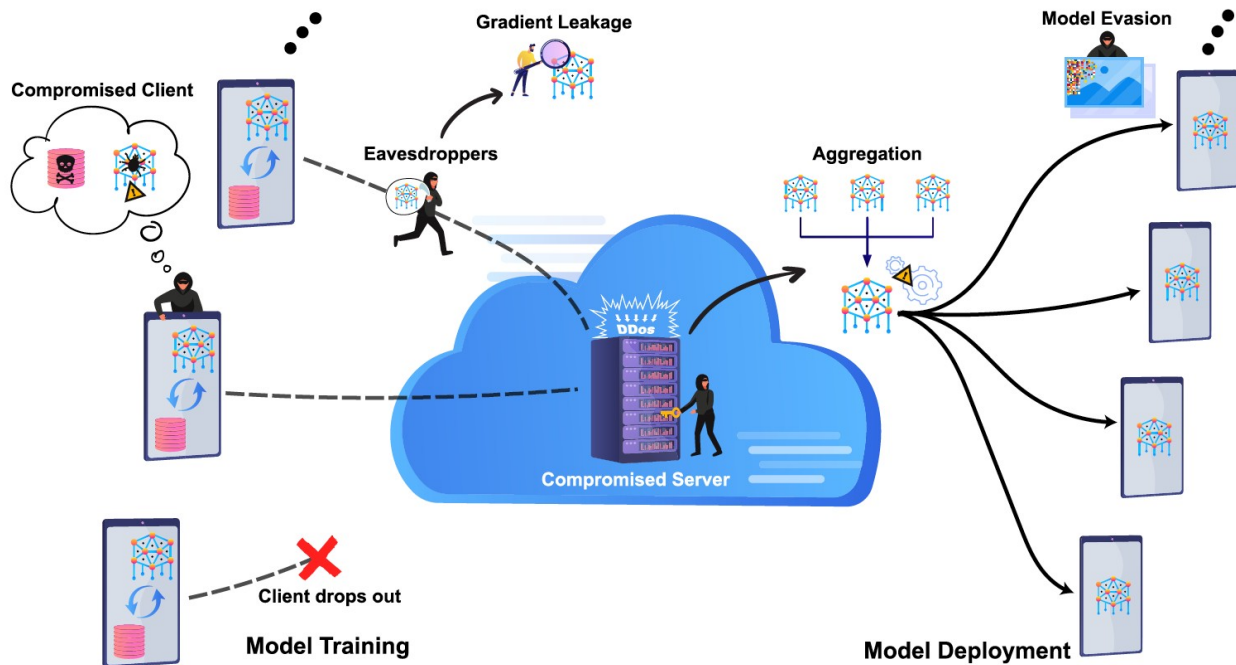


FIGURE 1. The lifecycle of FL process and the various sources of vulnerabilities.

From Bouacida et al. IEEE Access
2021

Fed-BioMed:

- VPN deployment for sandboxed environment
- Approval of training plan code
- Secure, encrypted communication
- Secure aggregation
- Differential privacy
- Code robustness and hardness

Collaborations

- **MAGNET (INRIA Lille):** focused on Decentralized / Federated Learning
- **Accenture:** consulting company interested in Federated Learning applied to healthcare
- **Unicancer consortium:** gathers 10 french hospitals that would adopt Fed-BioMed framework in the long run. Currently Fed-BioMed is implemented in CHU Rouen, Lacassagne Nice, and soon Curie Institute, Paris



Medical Data in Fed-BioMed

- Standard data:

- Image datasets
- CSV datasets

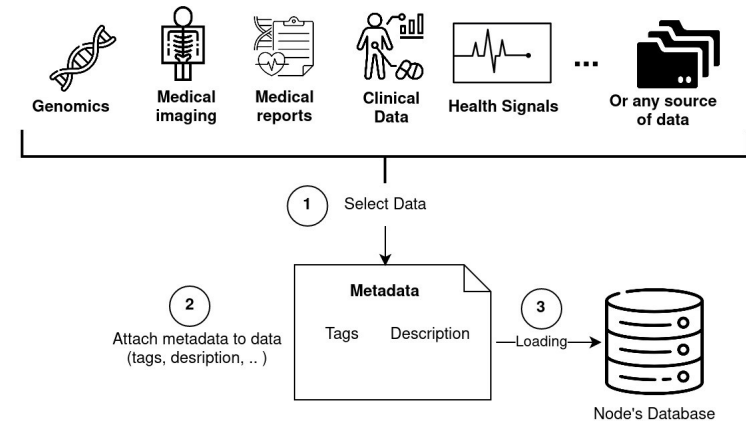
- Medical datasets:

- NIFTII datasets

- nifti_dataset_root_folder
- control_group
- subject_1.nii
- subject_2.nii
- ...
- disease_group
- subject_3.nii
- subject_4.nii
- ...

- BIDS datasets

- MedicalFolder_root/
- demographics.csv
- sub-01/
- T1/
- sub-01_xxx.nii.gz
- T2/
- sub-01_xxx.nii.gz



Additional features needed in Fed-BioMed

- Tool for displaying medical images
- Parse DICOM/PACS/... datasets
- Accessible GUI for loading medical dataset and monitoring experiment