# Federated Learning in Healthcare

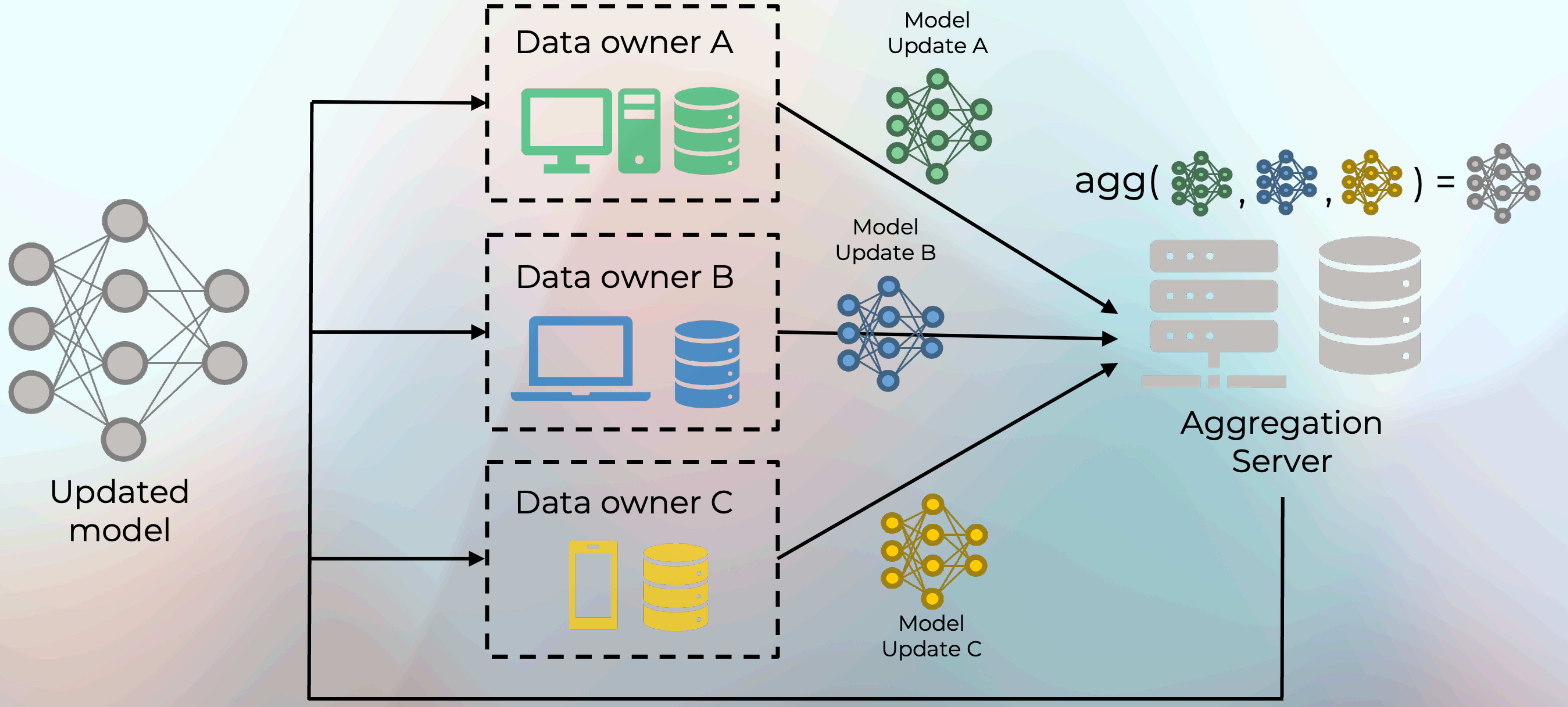
## Challenges and Research Directions

Mirko Polato
Assistant Professor @ Università degli Studi di Torino

FedMed workshop @ ICIAP - September 11, 2023

# Federated Learning in a nutshell

## Centralised FL



# Federated Learning & Healthcare

## An happy marriage

### Medical AI Needs Federated Learning, So Will Every Industry

Results published today in Nature Medicine demonstrate that federated learning builds powerful AI models that generalize across healthcare institutions, a finding that shows promise for further applications in energy, financial services, manufacturing and beyond.

September 15, 2021 by MONA FLORES

### Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data

Micah J. Sheller, Brandon Edwards, G. Anthony Reina, Jason Martin, Sarthak Pati, Aikaterini Kotrotsou, Mikhail Milchenko, Weilin Xu, Daniel Marcus, Rivka R. Colen & Spyridon Bakas

*Scientific Reports* **10**, Article number: 12598 (2020) | Cite this article

### Federated learning enables big data for rare cancer boundary detection

Sarthak Pati, Ujjwal Baid, Brandon Edwards, Micah Sheller, Shih-Han Wang, G. Anthony Reina, Patrick Foley, Alexey Gruzdev, Deepthi Karkada, Christos Davatzikos, Chiharu Sako, Satyam Ghodasara, Michel Bilello, Suyash Mohan, Philipp Vollmuth, Gianluca Brugnara, Chandrakanth J. Preetha, Felix Sahm, Klaus Maier-Hein, Maximilian Zenk, Martin Bendszus, Wolfgang Wick, Evan Calabrese, Jeffrey Rudie, … Spyridon Bakas + Show authors

*Nature Communications* **13**, Article number: 7346 (2022) | Cite this article

### Federated learning for predicting clinical outcomes in patients with COVID-19

Ittai Dayan, Holger R. Roth, Aoxiao Zhong, Ahmed Harouni, Amilcare Gentili, Anas Z. Abidin, Andrew Liu, Anthony Beardsworth Costa, Bradford J. Wood, Chien-Sung Tsai, Chih-Hung Wang, Chun-Nan Hsu, C. K. Lee, Peiying Ruan, Daguang Xu, Dufan Wu, Eddie Huang, Felipe Campos Kitamura, Griffin Lacey, Gustavo César de Antônio Corradi, Gustavo Nino, Hao-Hsin Shin, Hirofumi Obinata, Hui Ren, … Quanzheng Li + Show authors

*Nature Medicine* **27**, 1735–1743 (2021) | Cite this article

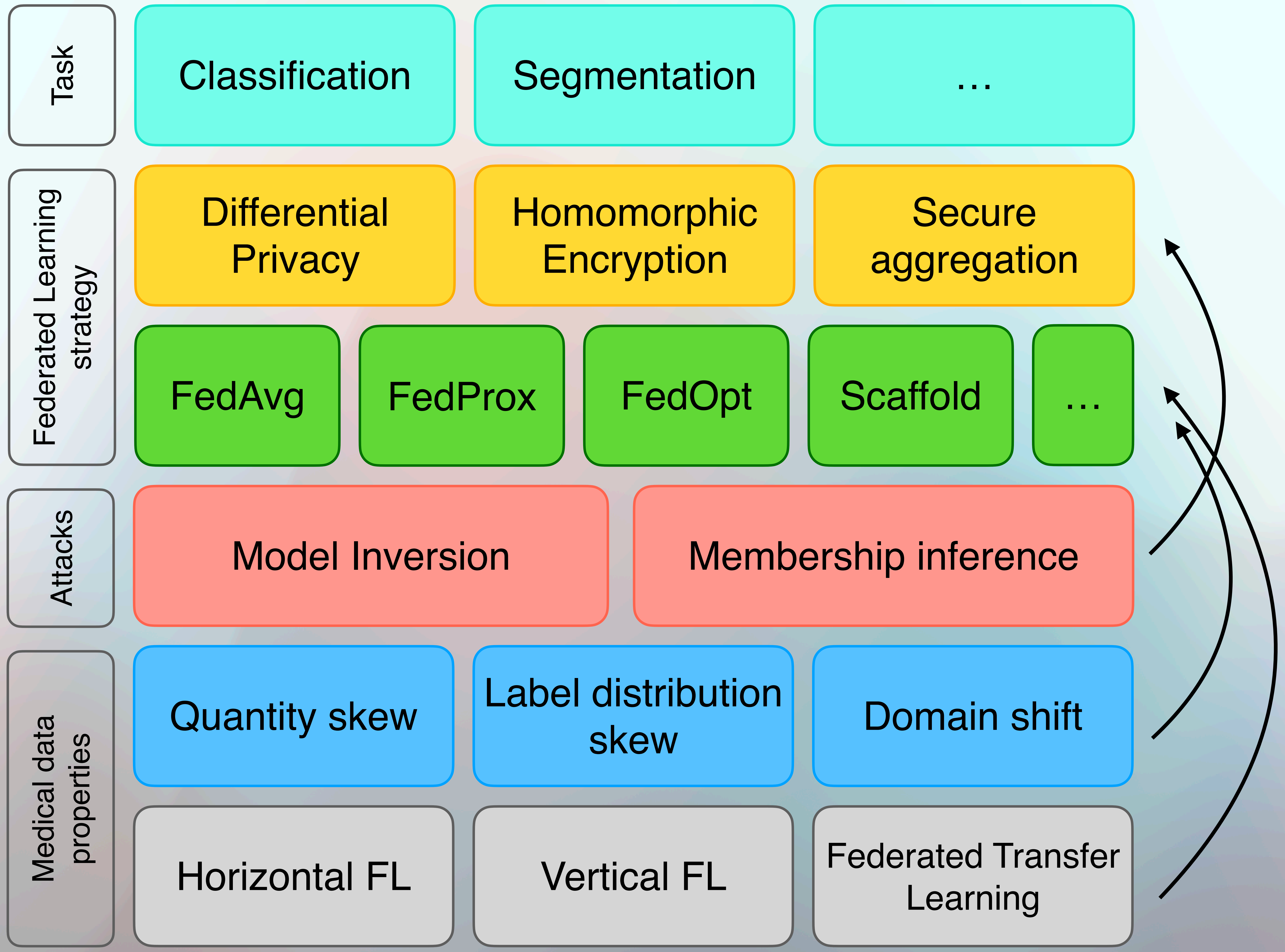
### The future of digital health with federated learning

Nicola Rieke, Jonny Hancox, Wenqi Li, Fausto Milletarì, Holger R. Roth, Shadi Albarqouni, Spyridon Bakas, Mathieu N. Galtier, Bennett A. Landman, Klaus Maier-Hein, Sébastien Ourselin, Micah Sheller, Ronald M. Summers, Andrew Trask, Daguang Xu, Maximilian Baust & M. Jorge Cardoso

*npj Digital Medicine* **3**, Article number: 119 (2020) | Cite this article

# Federated Learning in Healthcare

## The operational stack



# Typical FL setting in healthcare

## (Centralised) Cross-silo FL

Few institutions (<50)

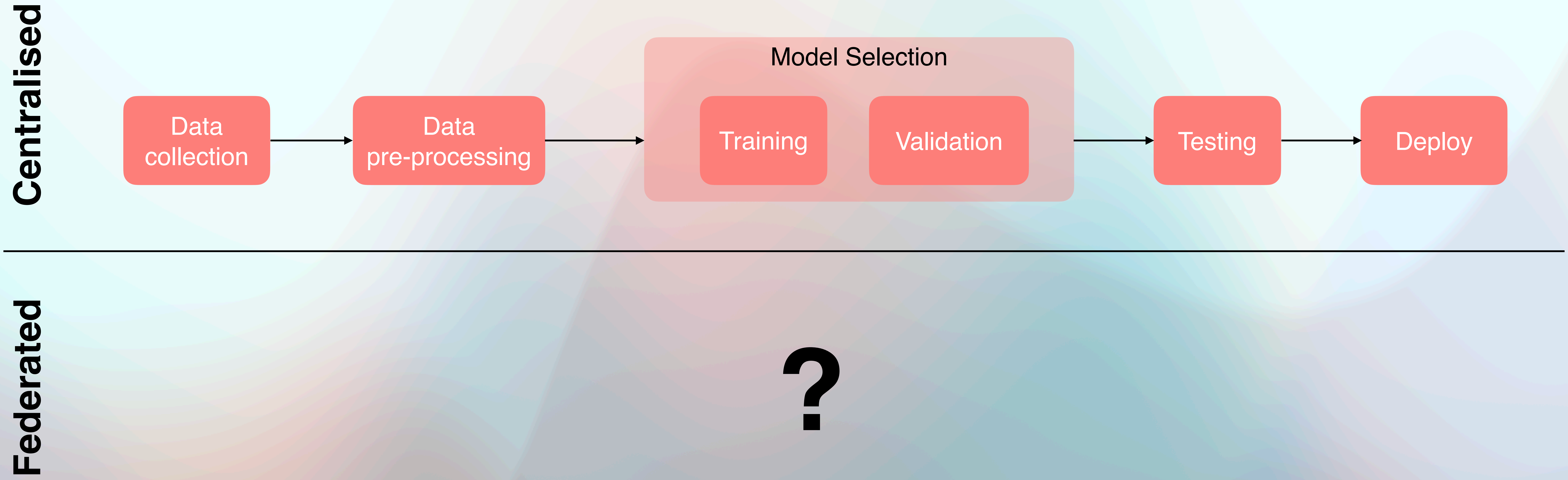
Stable connectivity

Relatively big datasets

Reliability

# Federated Learning in Healthcare

## From centralised to federated: Challenges & Research directions



# Non-iid data distribution

## Challenge



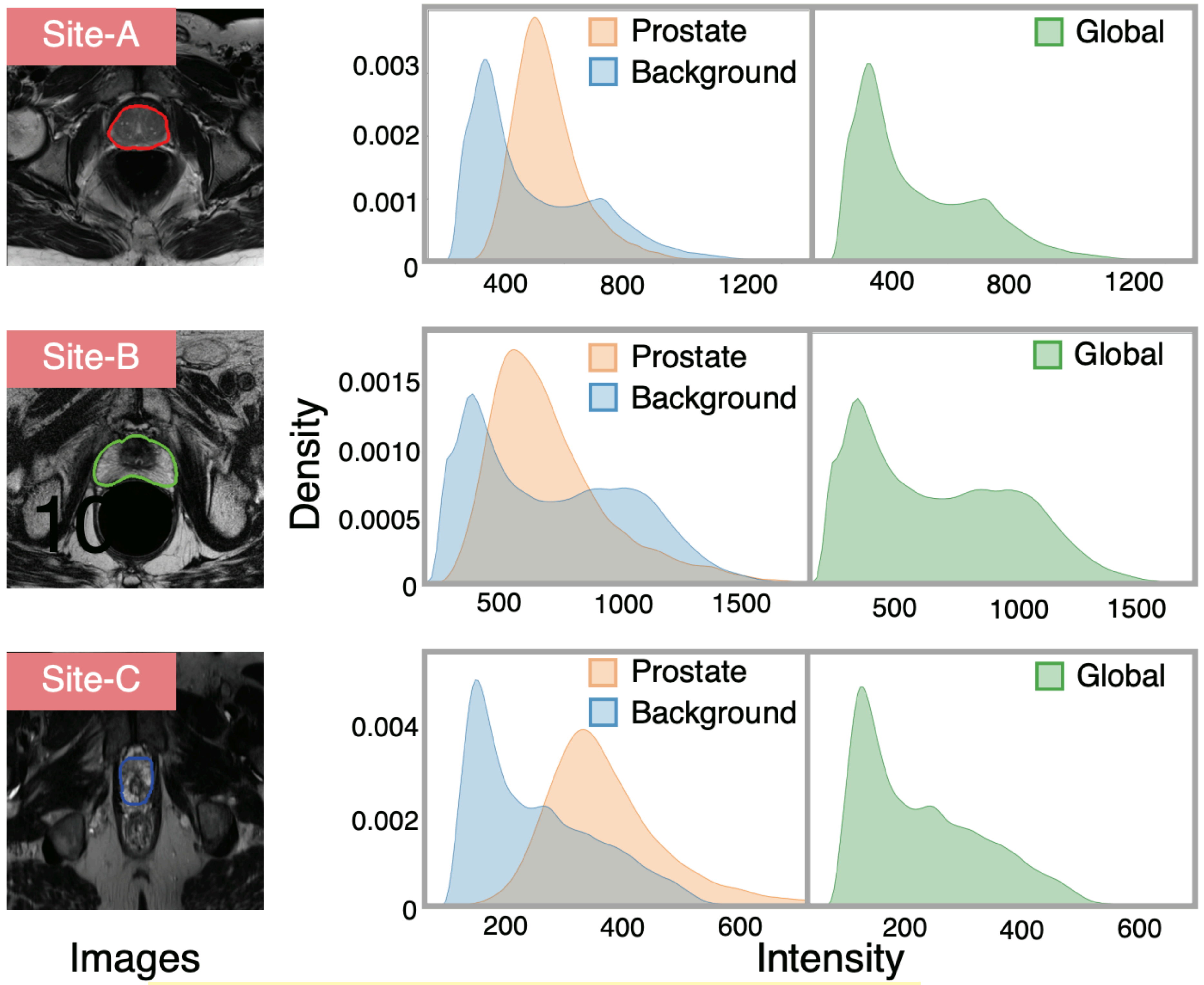
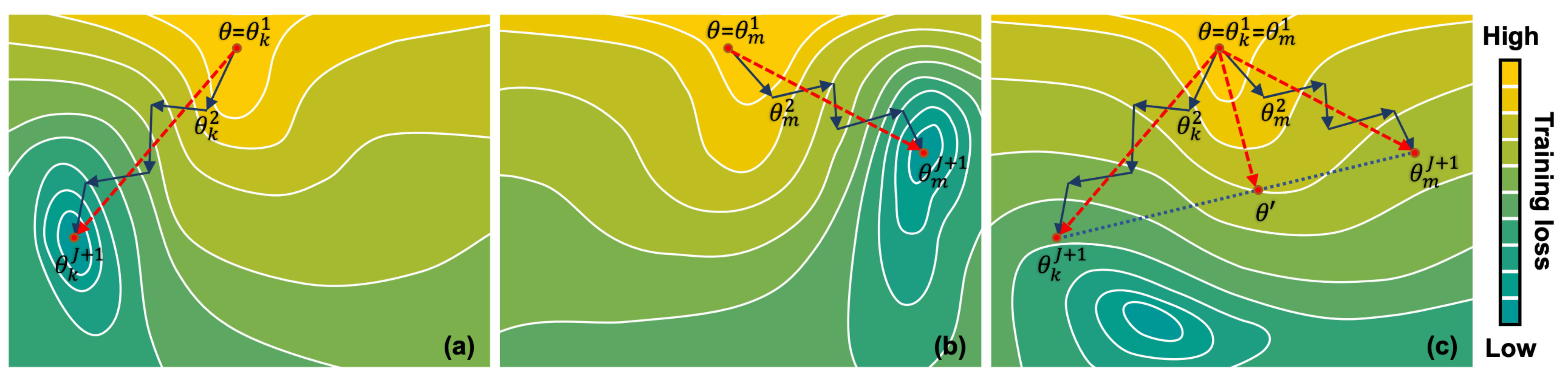



Figure 3: Domain shift among different medical sites. Region-wise and global intensity distribution of different sites for prostate MR images. Image courtesy to Xiao *et al.* (Xiao et al., 2022).



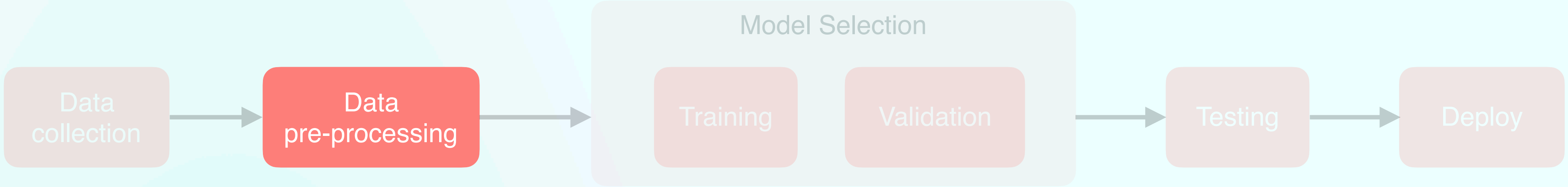

Xu, et al. Federated Cross Learning for Medical Image Segmentation. MIDL 2023

Hao Guan, Mingxia Liu. Federated Learning for Medical Image Analysis: A Survey. https://arxiv.org/abs/2306.05980

Pfitzner et al. Federated Learning in a Medical Context: A Systematic Literature Review. ACM Transactions on Internet Technology 2021

# Coordination needed

## Challenge



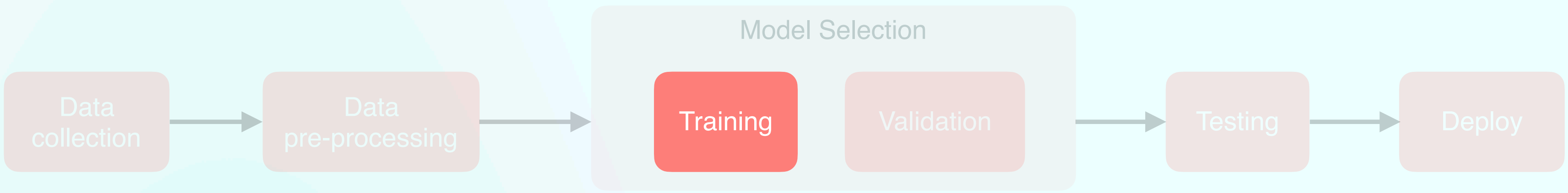

- Data must have the same format across institutions, e.g., *Fast Healthcare Interoperability Resources (FHIR)**
- Institutions **must coordinate** to agree upon which data pre-processing steps to perform
- In a Federated setting, some pre-processing operations may behave differently w.r.t. to the centralised setting, e.g., data standardization!

* https://www.hl7.org/fhir/

Hao Guan, Mingxia Liu. Federated Learning for Medical Image Analysis: A Survey. https://arxiv.org/abs/2306.05980

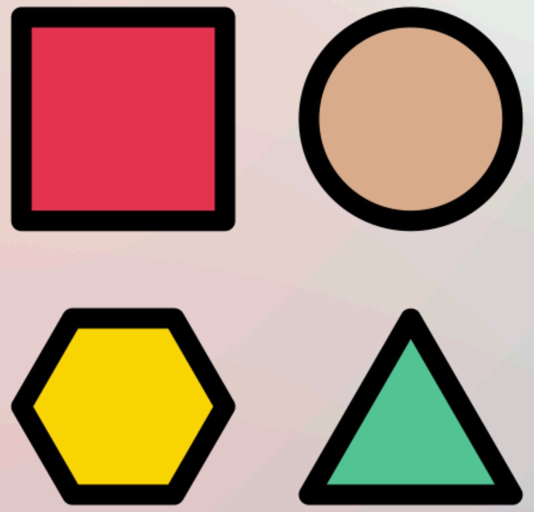
Pfitzner et al. Federated Learning in a Medical Context: A Systematic Literature Review. ACM Transactions on Internet Technology 2021

# Fast, Robust & Secure training




## Challenges

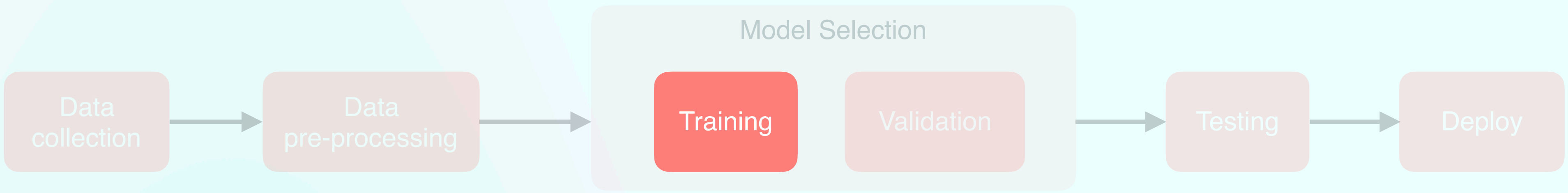
Fast



Robust to client heterogeneity

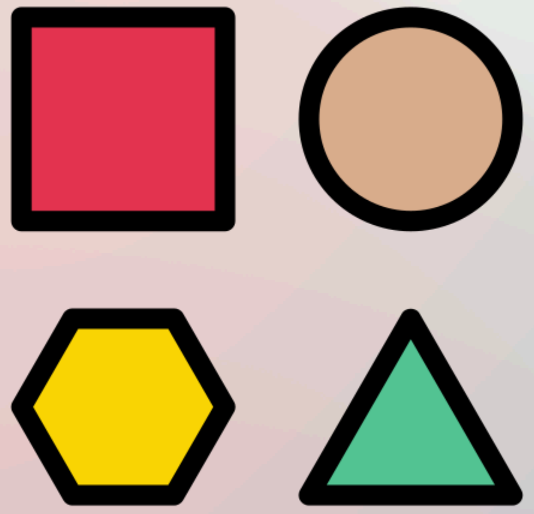
Privacy

# Fast, Robust & Secure training




## Challenges

Fast



Robust to client heterogeneity

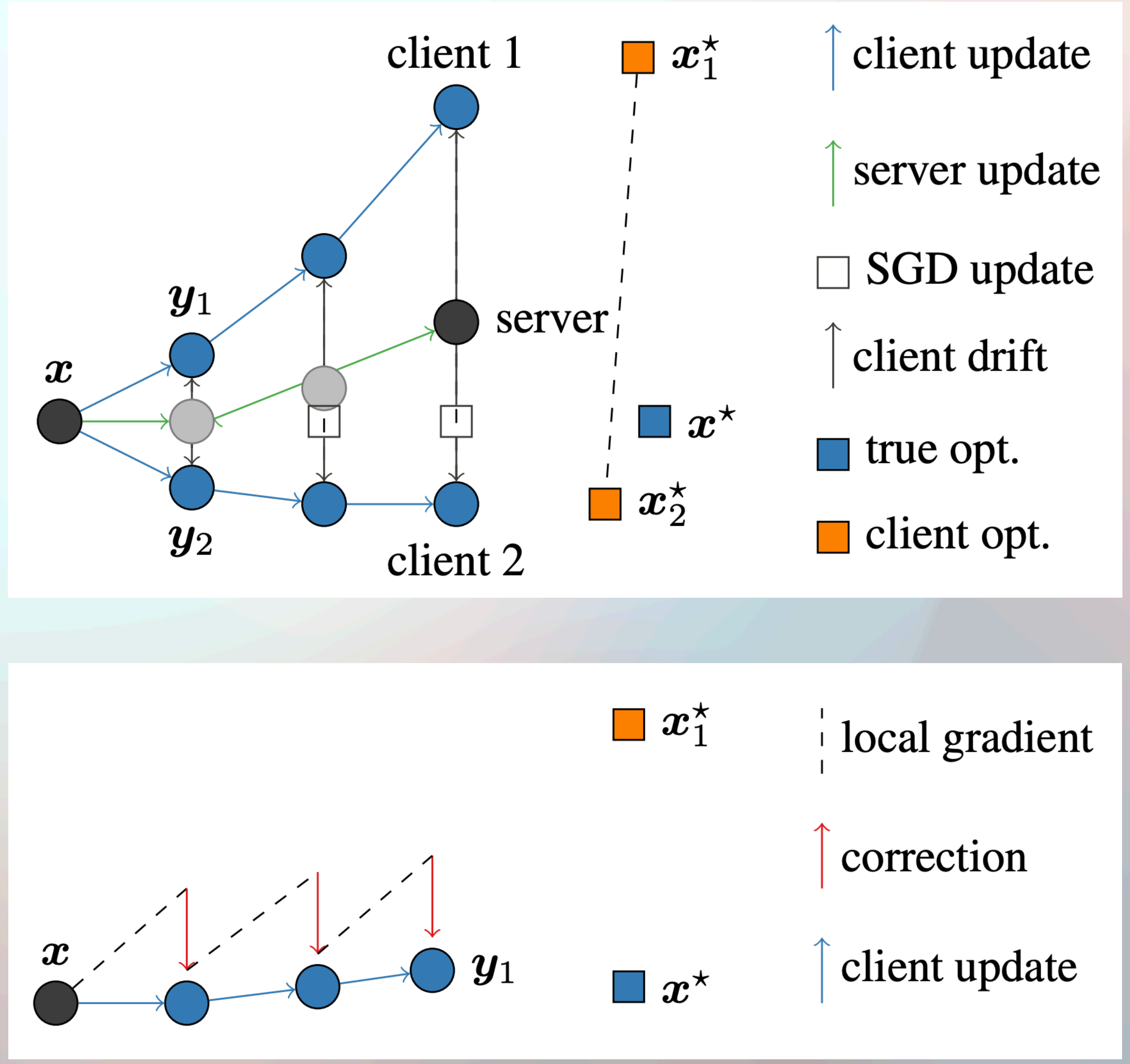
Privacy

# Handling non-iid clients

## A hot research direction in FL

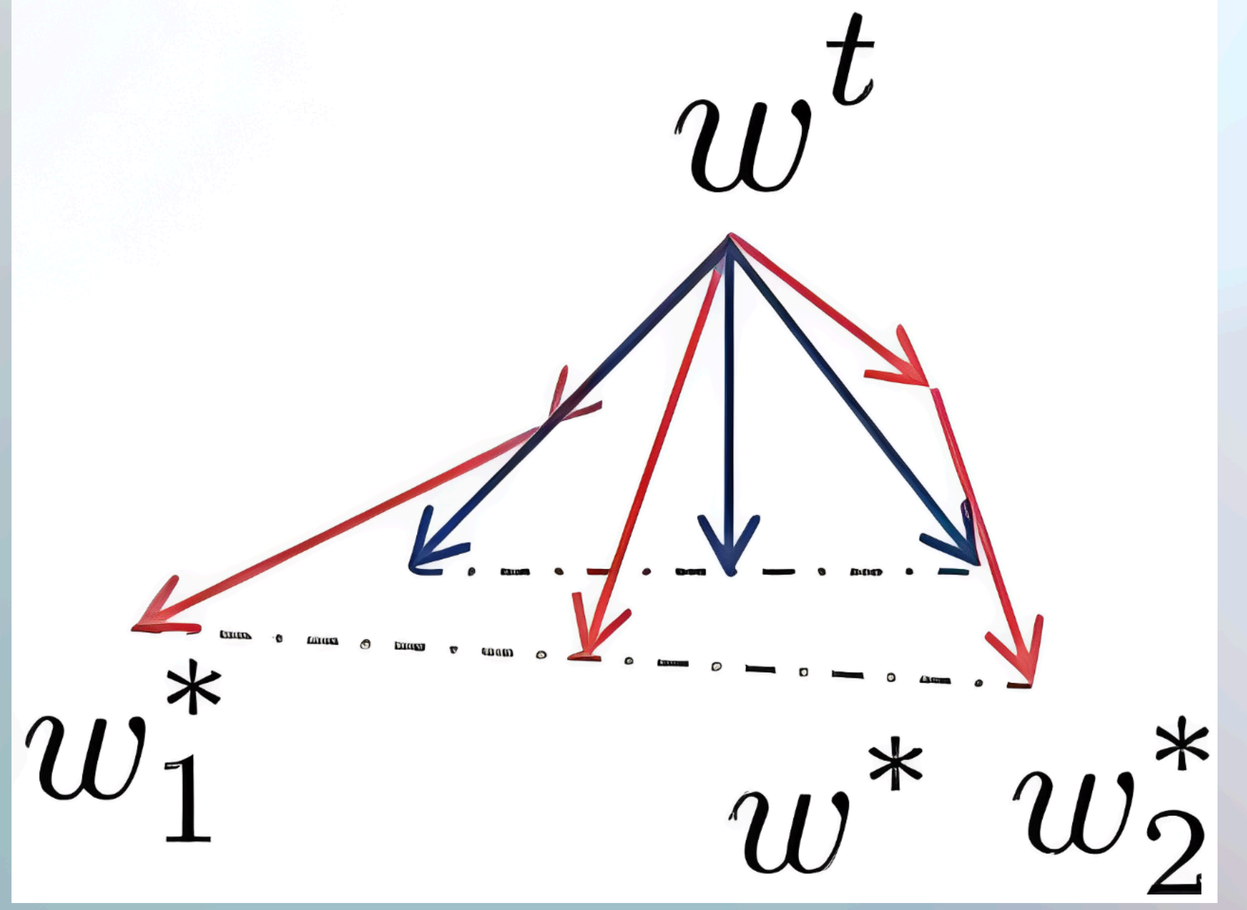


### SCAFFOLD




### FedProx

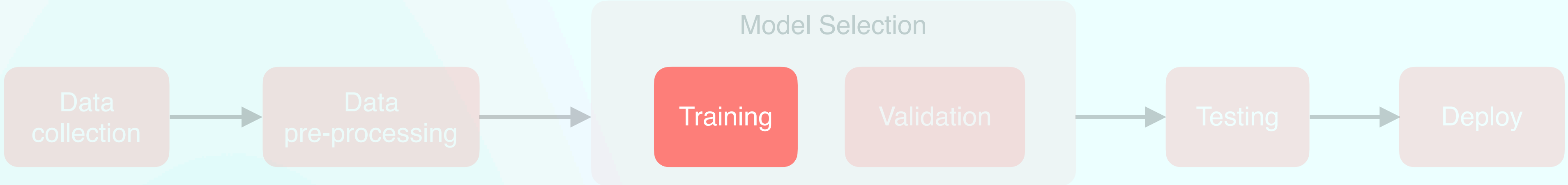
$$\min_{w_k} F_k(w_k) + \frac{\mu}{2}\left\|w_k - w^t\right\|^2$$

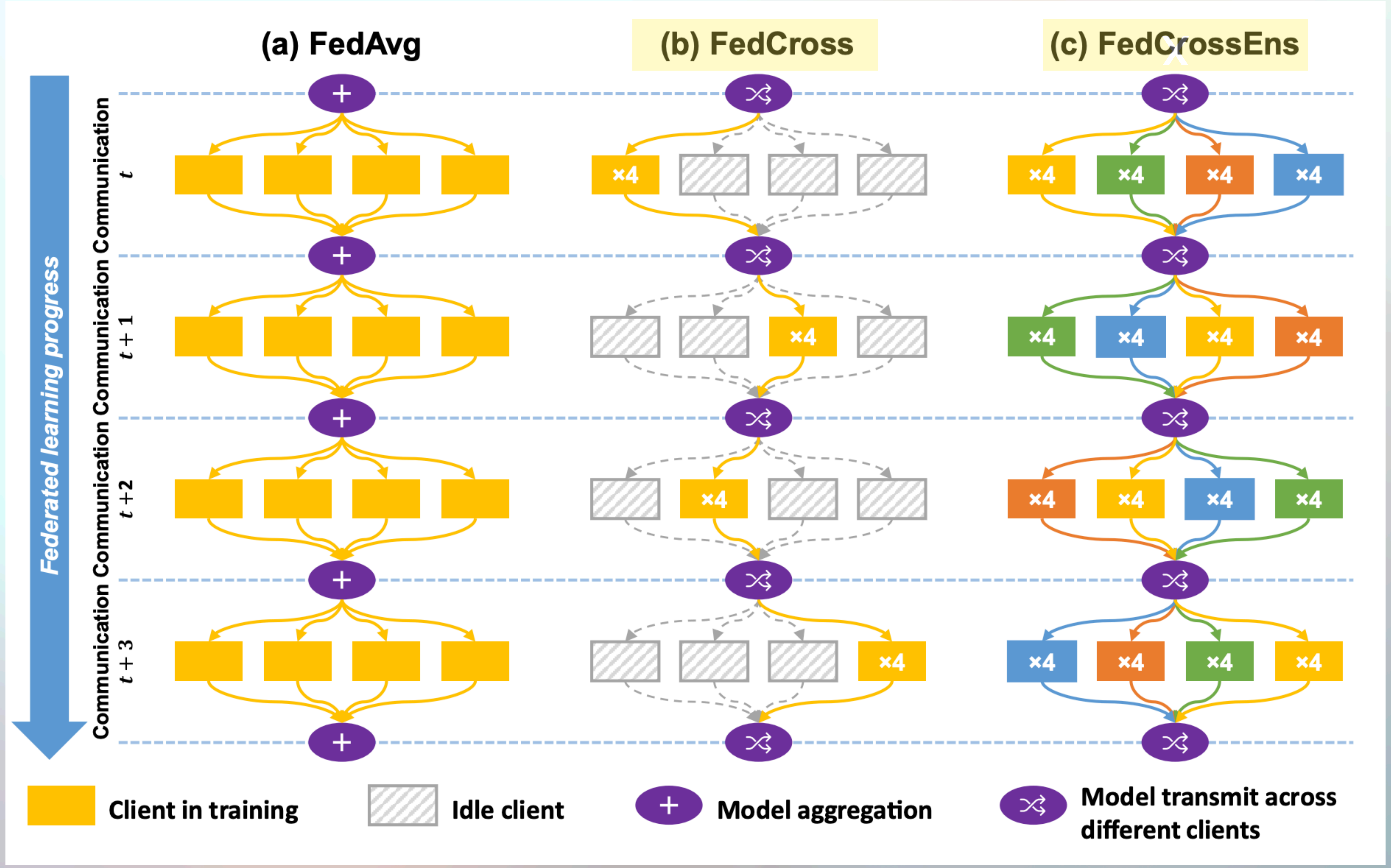


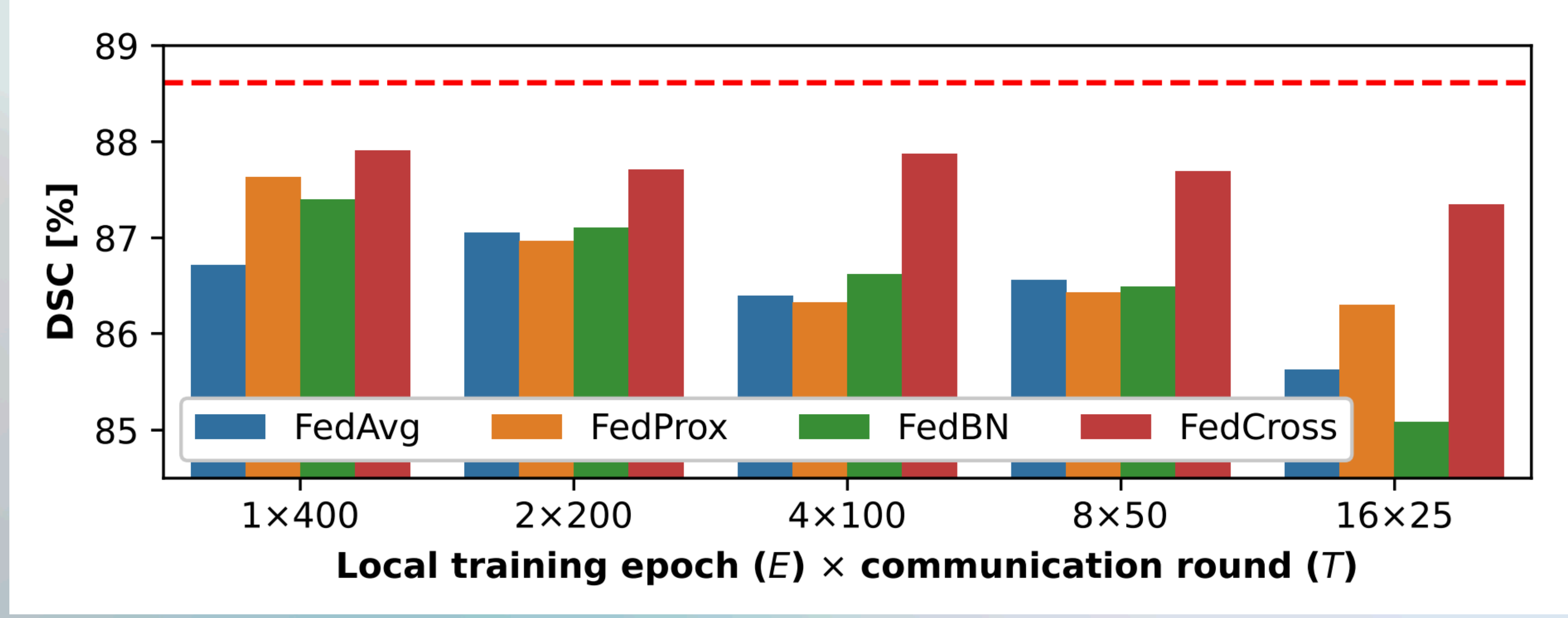

Karimireddy, et al. SCAFFOLD: Stochastic Controlled Averaging for Federated Learning. ICML 2020
Li, et al. Federated Optimization in Heterogeneous Networks. MLSys 2020

# Handling non-iid clients

## A hot research direction in FL



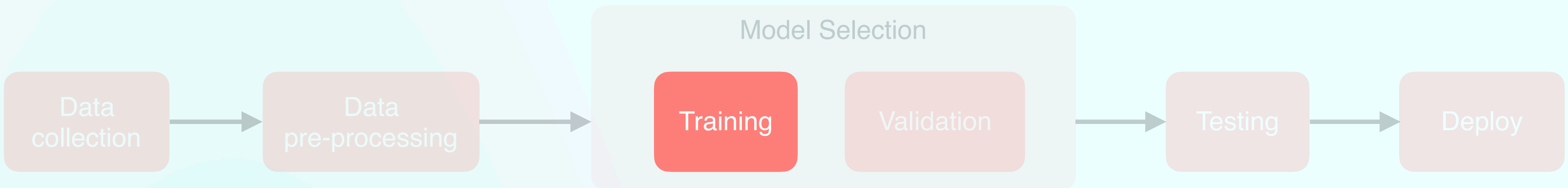



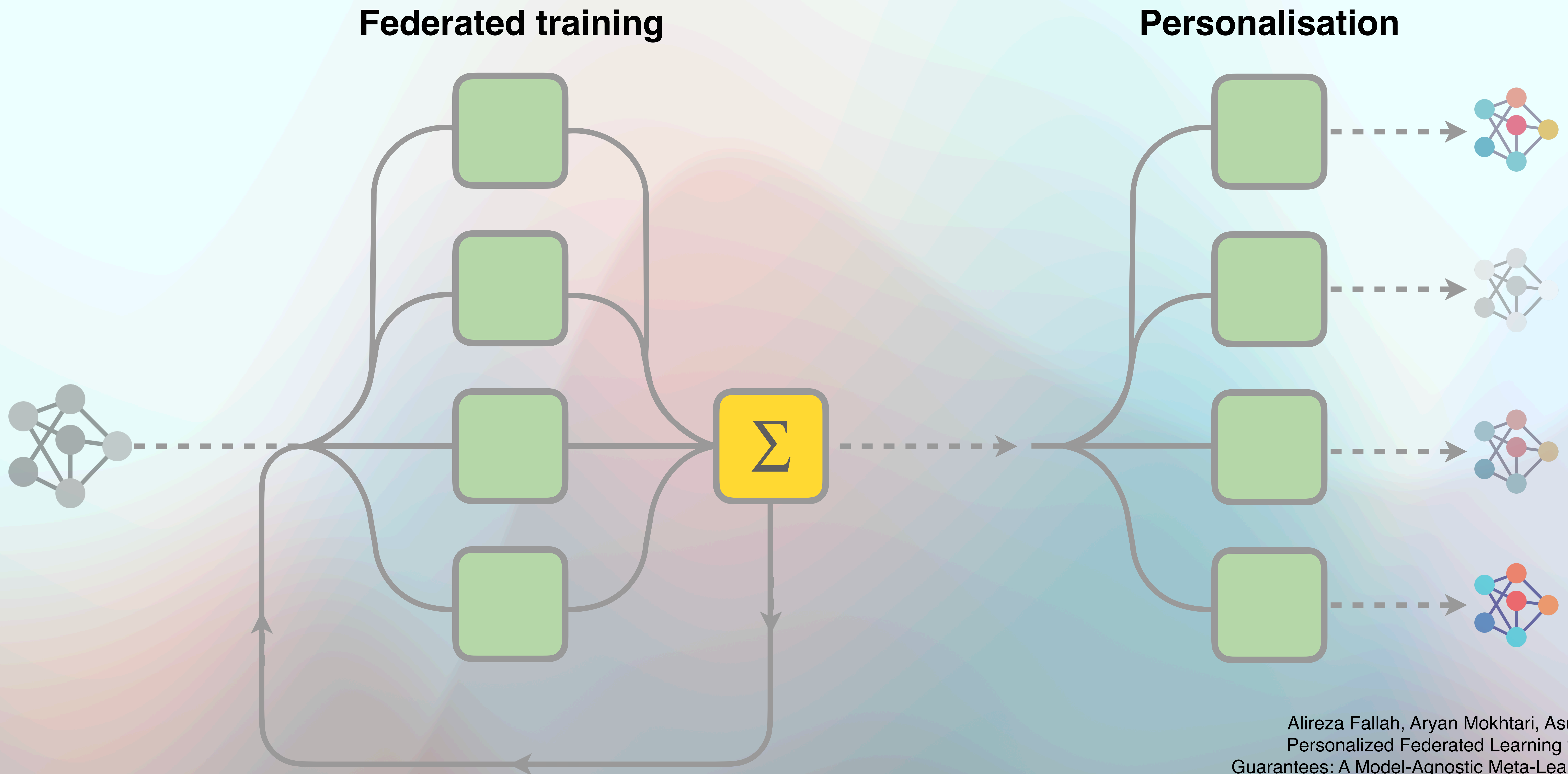




Xu, et al. Federated Cross Learning for Medical Image Segmentation. MIDL 2023

# Personalized FL

## Research direction



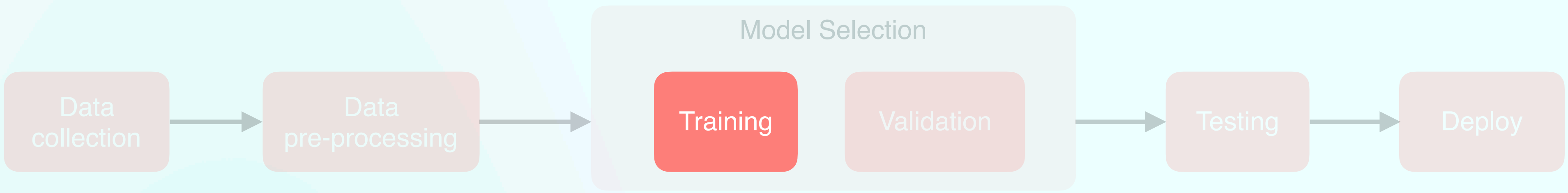





Alireza Fallah, Aryan Mokhtari, Asuman Ozdaglar. Personalized Federated Learning with Theoretical Guarantees: A Model-Agnostic Meta-Learning Approach. NeurIPS 2020

Hao Guan, Mingxia Liu. Federated Learning for Medical Image Analysis: A Survey. https://arxiv.org/abs/2306.05980

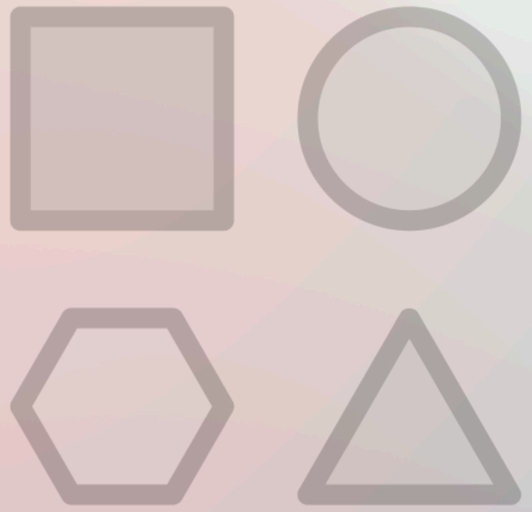
# Fast, Robust & Secure training




## Challenges



Fast

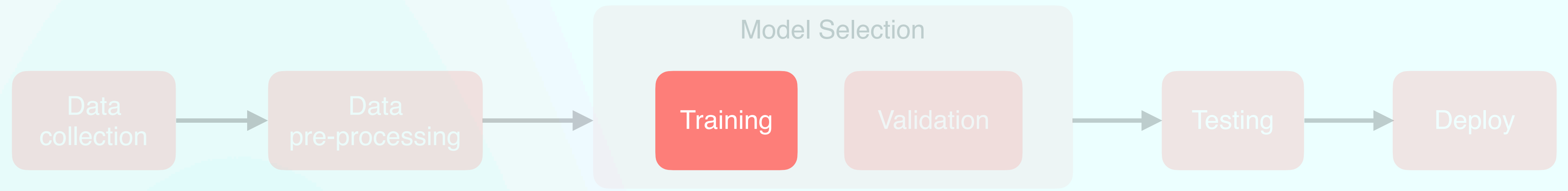


Robust to client heterogeneity

Privacy

# Differential Privacy

## Research direction




### DPFed-Post

$N$ total clients, local mini-batch size $B$, local epochs $E$, communication rounds $T_{cl}$, learning rate $\eta$, sensitivity $S$ and post-processing parameter $P$.

---

Initialize $\mathbf{w}_0$ and send the model to clients

**for** $r = 1, \ldots T_{cl}$

  Select $K$ clients randomly

  **for** each selected client $k = 1, \ldots, K$

    $\mathbf{w}_k^r \leftarrow \text{ClientUpdate}(k, \mathbf{w}^{r-1})$

    $\Delta\mathbf{w}_k^r \leftarrow \mathbf{w}_k^r - \mathbf{w}^{r-1}$

    $\Delta\hat{\mathbf{w}}_k^r \leftarrow \Delta\mathbf{w}_k^r / \max\left(1, \frac{||\Delta\mathbf{w}_k^r||_2}{S}\right)$

  $\Delta\mathbf{w}^r \leftarrow \frac{\sum_{k=1}^{K} \Delta\hat{\mathbf{w}}_k^r + \mathcal{G}(0, S\sigma\mathbf{I})}{K}$

  $\Delta\hat{\mathbf{w}}^r \leftarrow \Delta\mathbf{w}^r / \max\left(1, \frac{||\Delta\mathbf{w}^r||_2}{P}\right)$

  $\mathbf{w}^r \leftarrow \mathbf{w}^{r-1} + \Delta\hat{\mathbf{w}}^r$

**ClientUpdate**$(k, \mathbf{w})$

  for client $k$

  **for** $i = 1, \ldots, E$

    **for** local batches $b$

      $\mathbf{w} \leftarrow \mathbf{w} - \eta\nabla l(b; \mathbf{w})$

  **return** $\mathbf{w}$ to server

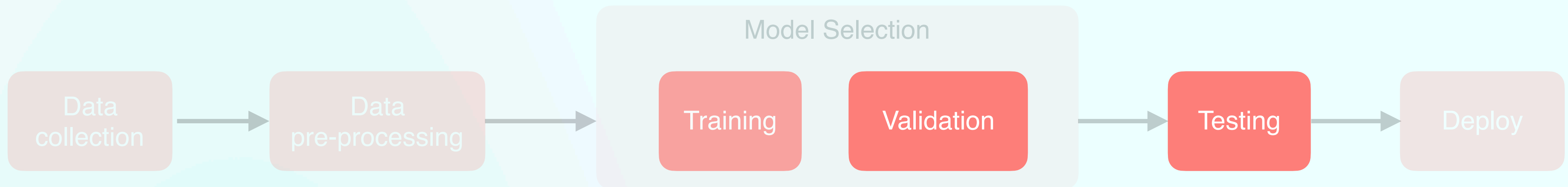
| Metric | Model | Non-Private Centralized | Non-Private StdFed |
|---|---|---|---|
| C-index ↑ | DeepHit | $0.66 \pm 0.02$ | $0.67 \pm 0.02$ |
| | CoxPH | $0.66 \pm 0.01$ | $0.67 \pm 0.03$ |
| | CoxCC | $0.63 \pm 0.02$ | $0.68 \pm 0.01$ |
| | CoxTime | $0.64 \pm 0.01$ | $0.67 \pm 0.01$ |

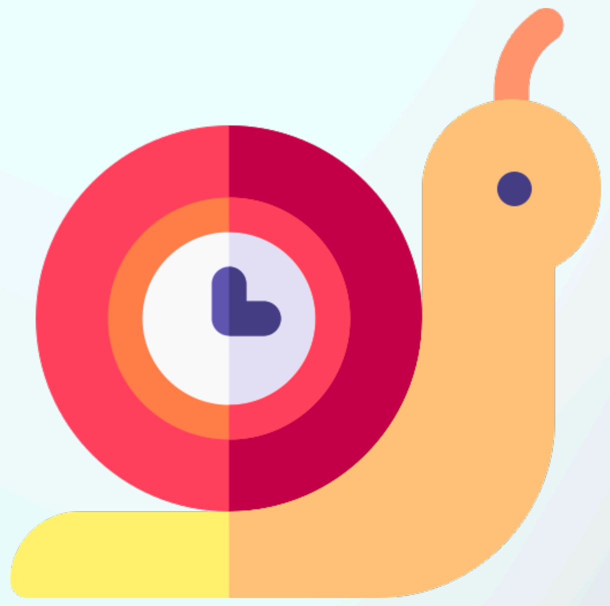
| Metric | Model | $(\epsilon = 5.4, \delta = 10^{-3})$ DPFed | $(\epsilon = 5.4, \delta = 10^{-3})$ $\text{DPFed}_{\text{post}}$ | $(\epsilon = 8.9, \delta = 10^{-3})$ DPFed | $(\epsilon = 8.9, \delta = 10^{-3})$ $\text{DPFed}_{\text{post}}$ |
|---|---|---|---|---|---|
| C-index ↑ | DeepHit | $0.47 \pm 0.03$ | $\mathbf{0.56 \pm 0.04}$ | $0.54 \pm 0.03$ | $\mathbf{0.59 \pm 0.03}$ |
| | CoxPH | $0.45 \pm 0.70$ | $\mathbf{0.62 \pm 0.02}$ | $0.47 \pm 0.05$ | $\mathbf{0.64 \pm 0.03}$ |
| | CoxCC | $0.58 \pm 0.05$ | $\mathbf{0.61 \pm 0.02}$ | $0.62 \pm 0.02$ | $\mathbf{0.64 \pm 0.03}$ |
| | CoxTime | $0.57 \pm 0.07$ | $\mathbf{0.62 \pm 0.02}$ | $0.61 \pm 0.03$ | $\mathbf{0.63 \pm 0.02}$ |

Shadi Rahimian, et al. Practical Challenges in Differentially-Private Federated Survival Analysis of Medical Data. CHIL 2022

# Validating (& Testing)

## Challenge



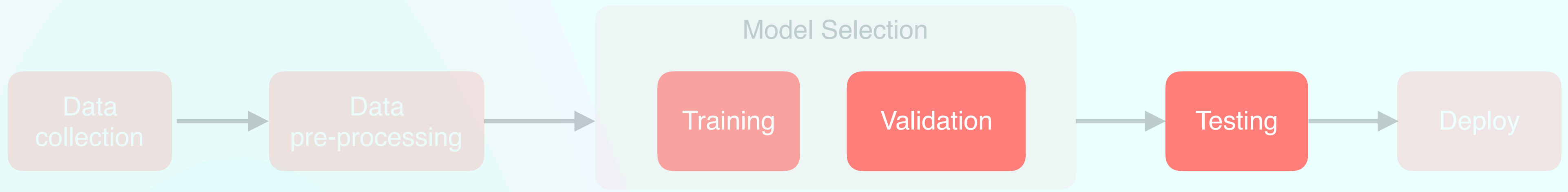



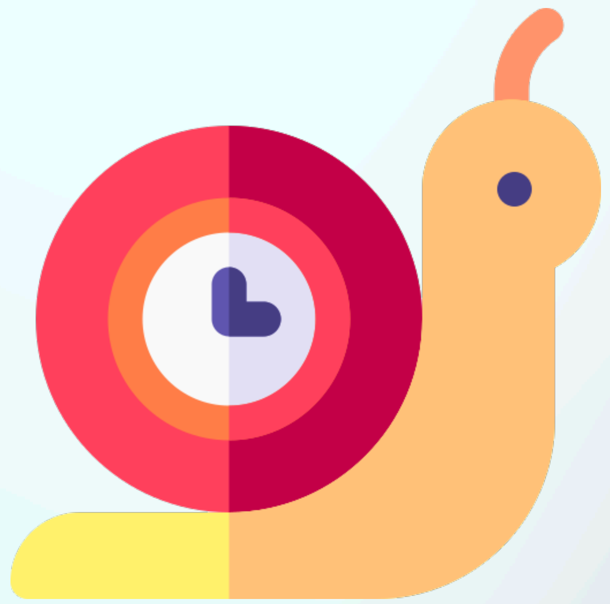
Slow process

Require a separate test set for each client

Benchmark

# Validating (& Testing)
## Challenge



Slow process

Require a separate test set for each client



Benchmark

# Neural Architecture Search

## A promising direction for federated model selection


Hieu Pham, et al. Efficient Neural Architecture Search via Parameter Sharing. ICML 2018

Chaoyang He, Haishan Ye, Li Shen, and Tong Zhang. Milenas: Efficient neural architecture search via mixed-level reformulation. CVPR 2020

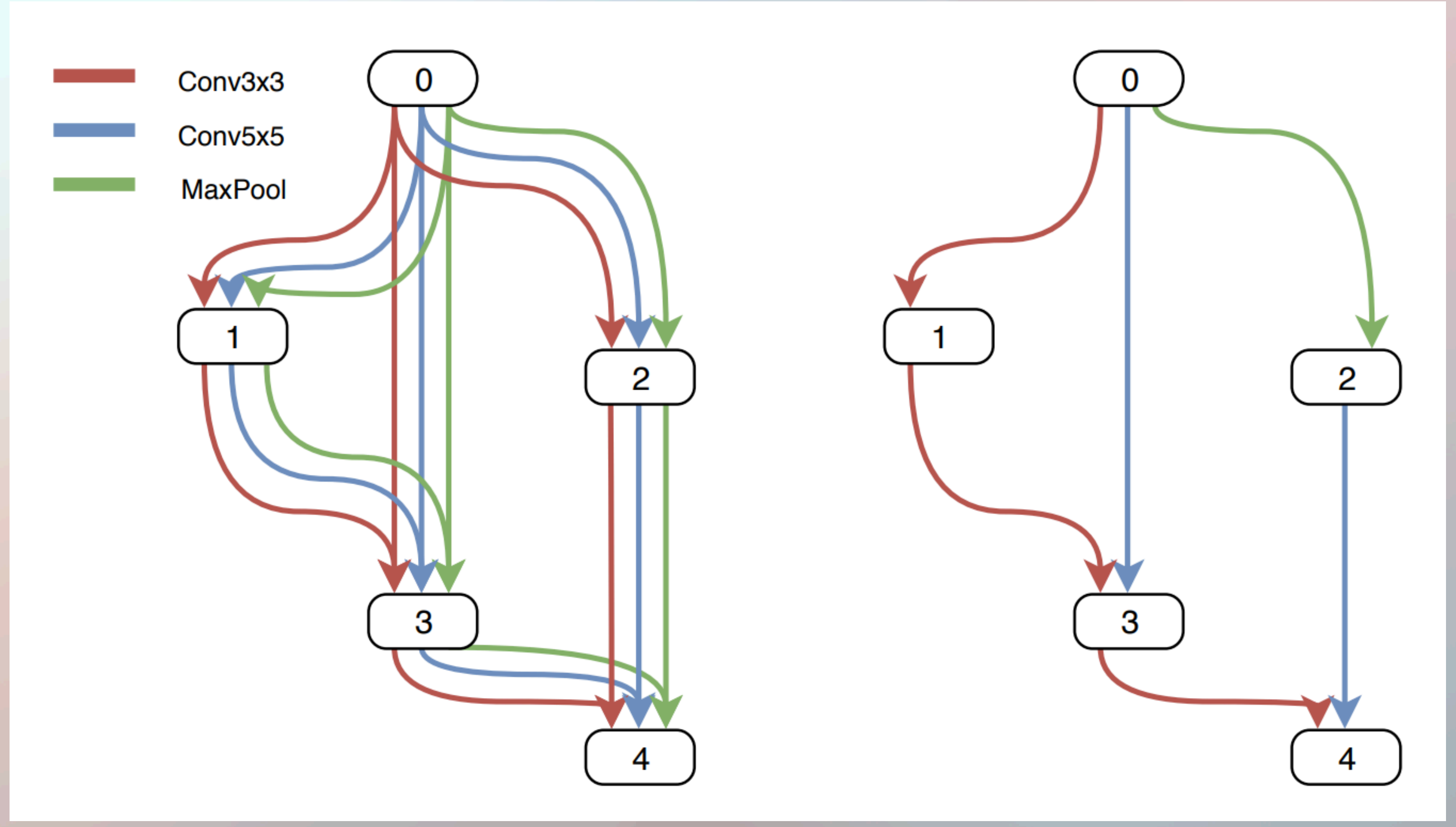
Hanxiao Liu, Karen Simonyan, Yiming Yang. DARTS: Differentiable Architecture Search. ICLR 2019

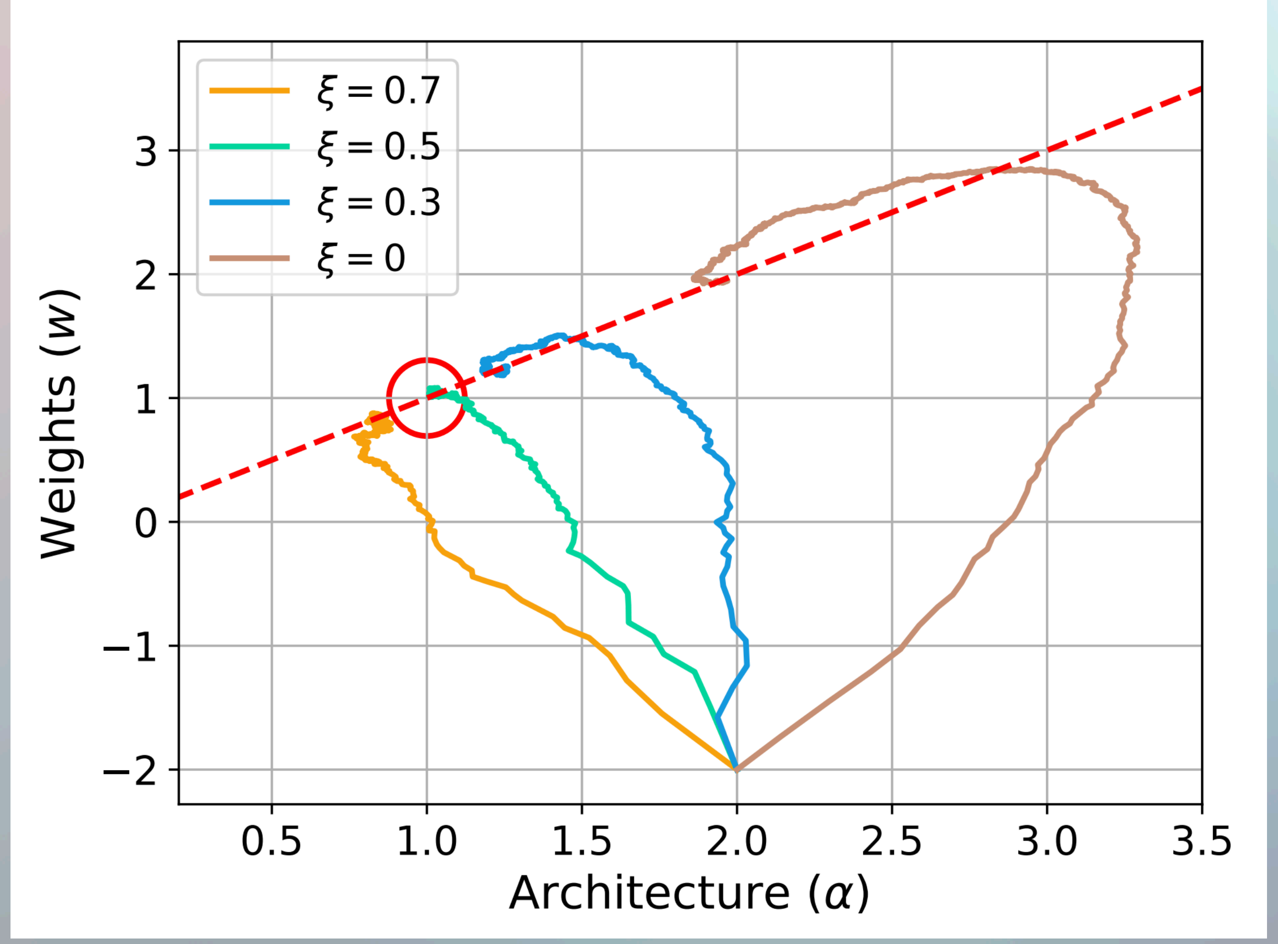

$$\min_{\alpha} L_{\text{valid}}(\mathbf{w}, \alpha) \quad \text{s.t.} \quad \mathbf{w} \in \arg\min_{\mathbf{u}} L_{\text{train}}(\mathbf{u}, \alpha)$$

"Defines" an architecture

$$\min_{\alpha, \mathbf{w}} L_{\text{train}}(\mathbf{w}, \alpha) + L_{\text{valid}}(\mathbf{w}, \alpha)$$

Possibly weighted by an hyper-parameter $\lambda$



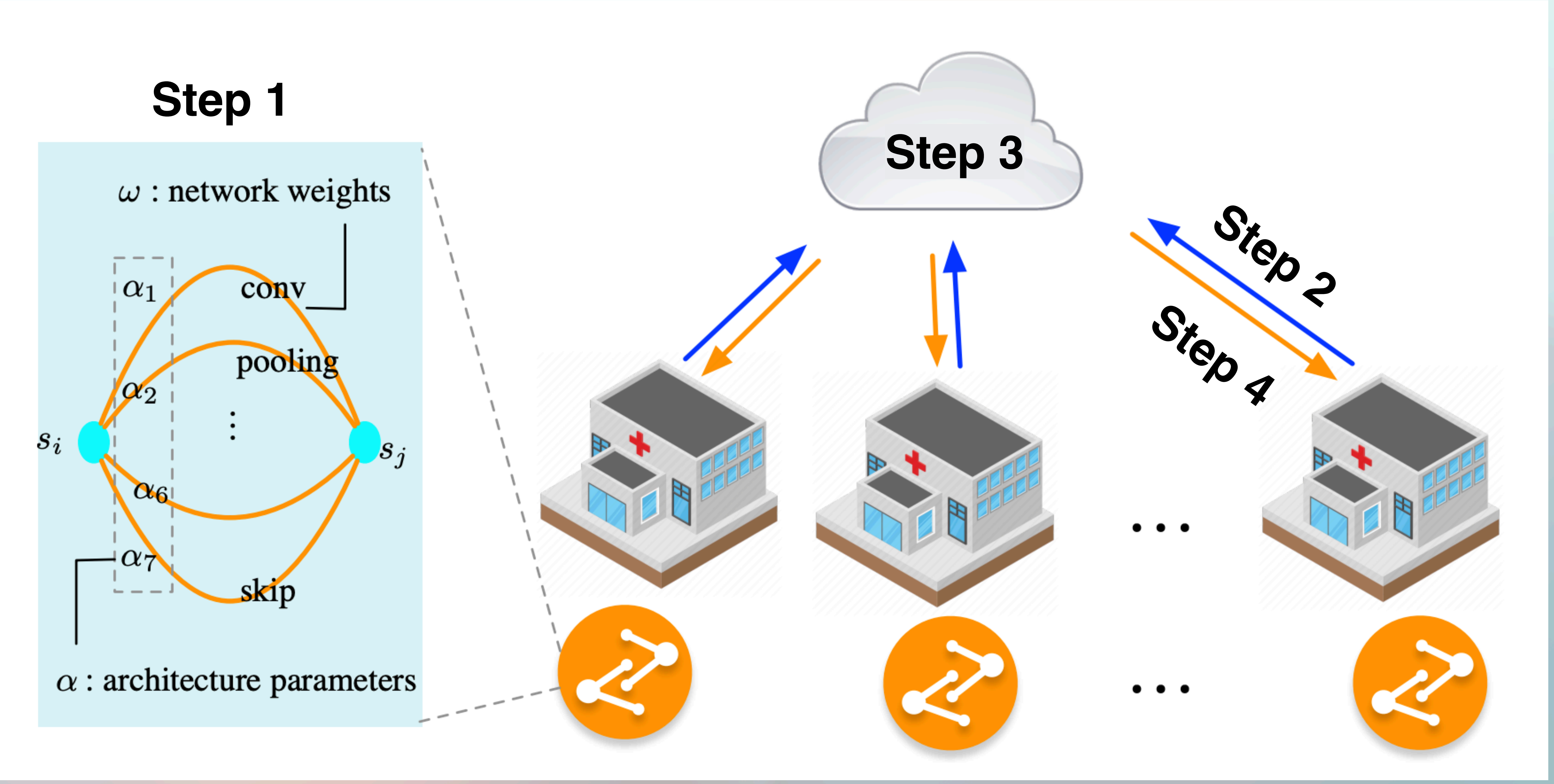



# FedNAS

## Federated application of gradient-based NAS

He, et al. FedNAS: Federated Deep Learning via Neural Architecture Search. CVPR 2020 Workshop on Neural Architecture Search and Beyond for Representation Learning

He, et al. MiLeNAS: Efficient Neural Architecture Search via Mixed-Level Reformulation. CVPR 2020

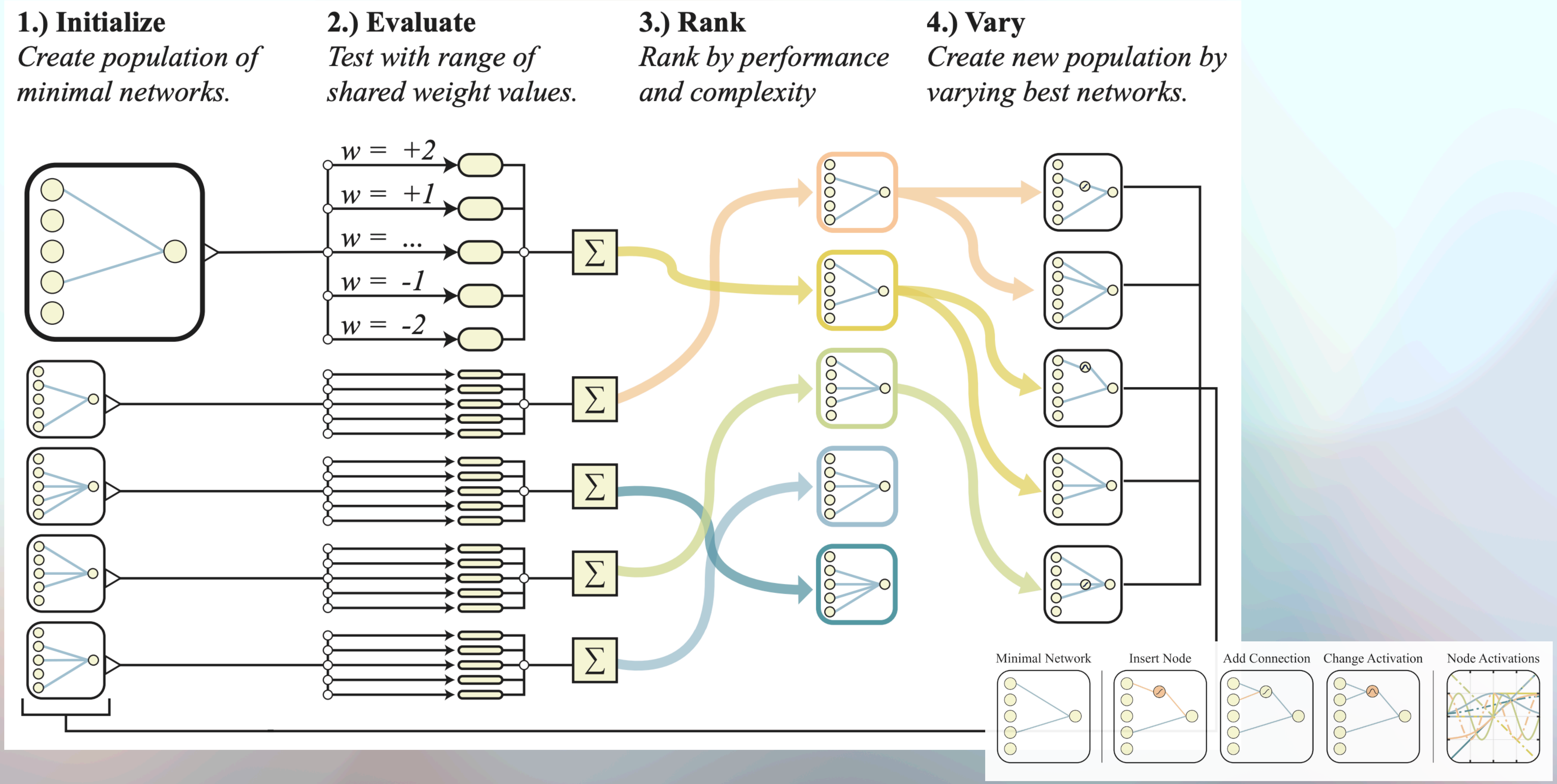



1. Local search (**architecture** + parameters), via stochastic gradient descent
2. Clients send the gradients to the server for both architectural parameters and network parameters
3. Server merges the gradients
4. Server sends the updated parameters to the clients

# Weight Agnostic NAS

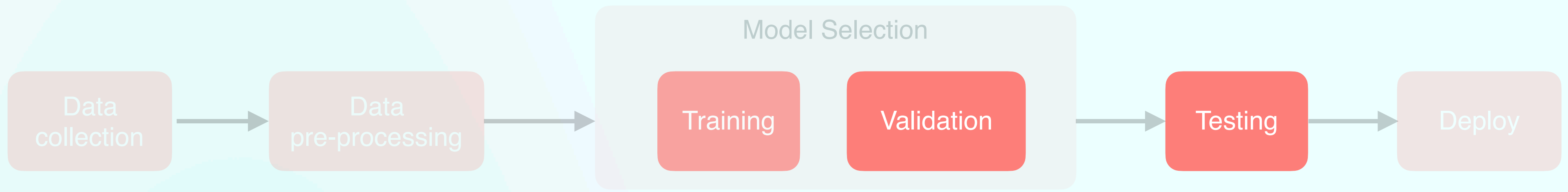
## A promising direction to speed up federated NAS

Adam Gaier, David Ha. Weight Agnostic Neural Networks. NeurIPS 2019



# Validating (& Testing)

## Challenge






Slow process

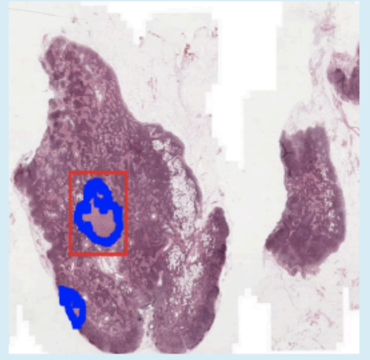
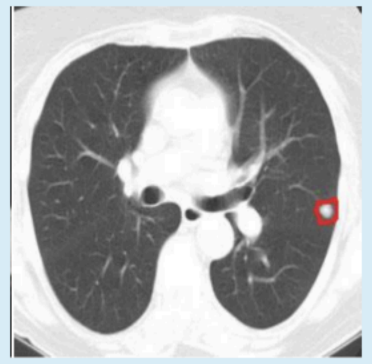
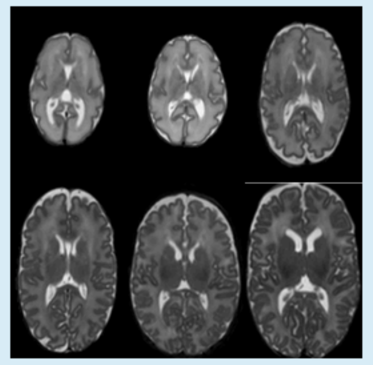
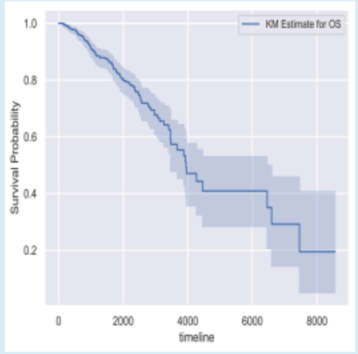
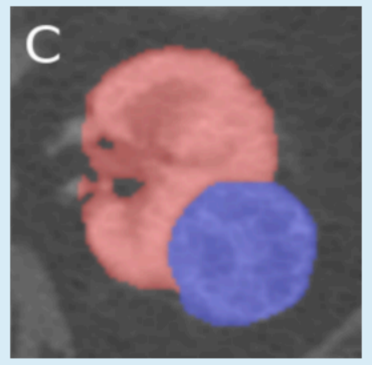
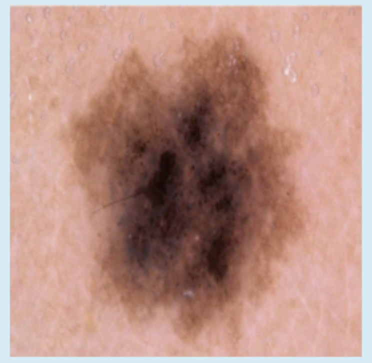
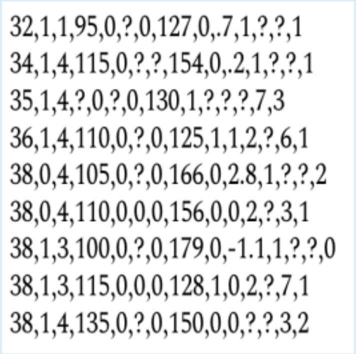
Require a separate test set for each client

Benchmark

# FLamby

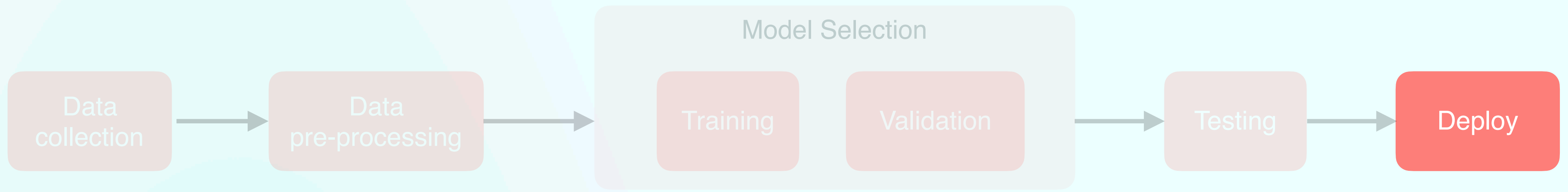
## Federated Learning AMple Benchmark of Your cross-silo strategies

Du Terrail et al. FLamby: Datasets and Benchmarks for Cross-Silo Federated Learning in Realistic Healthcare Settings. NeurIPS 2022 (Track on Datasets and Benchmarks)

| Dataset | Fed-Camelyon16 | Fed-LIDC-IDRI | Fed-IXI | Fed-TCGA-BRCA | Fed-KITS2019 | Fed-ISIC2019 | Fed-Heart-Disease |
|---|---|---|---|---|---|---|---|
| **Input (x)** | Slides | CT-scans | T1WI | Patient info. | CT-scans | Dermoscopy | Patient info. |
| **Preprocessing** | Matter extraction + tiling | Patch Sampling | Registration | None | Patch Sampling | Various image transforms | Removing missing data |
| **Task type** | binary classification | 3D segmentation | 3D segmentation | survival | 3D segmentation | multi-class classification | binary classification |
| **Prediction (y)** | Tumor on slide | Lung Nodule Mask | Brain mask | Risk of death | Kidney and tumor masks | Melanoma class | Heart disease |
| **Center extraction** | Hospital | Scanner Manufacturer | Hospital | Group of Hospitals | Group of Hospitals | Hospital | Hospital |
| **Thumbnails** |  |  |  |  |  |  |  |
| **Original paper** | Litjens *et al.* 2018 | Armato *et al.* 2011 | Perez *et al.* 2021 | Liu *et al.* 2018 | Heller *et al.* 2019 | Tschandl *et al.* 2018 / Codella *et al.* 2017 / Combalia *et al.* 2019 | Janosi *et al.* 1988 |
| **# clients** | 2 | 5 | 3 | 5 | 6 | 5 | 4 |
| **# examples** | 399 | 1,018 | 566 | 1, 088 | 96 | 23, 247 | 740 |
| **# examples per center** | 239, 150 | 670, 205, 69, 74 | 311, 181, 74 | 311, 196, 206, 162 51 | 12, 14, 12, 12, 16, 30 | 12413, 3954, 3363, 225 819, 439 | 303, 261, 46, 130 |
| **Model** | DeepMIL [66] | Vnet [100, 102] | 3D U-net [25] | Cox Model [33] | nnU-Net [69] | efficientnet [119] + linear layer | Logistic Regression |
| **Metric** | AUC | DICE | DICE | C-index | DICE | Balanced Accuracy | Accuracy |
| **Size** | 50G (850G total) | 115G | 444M | 115K | 54G | 9G | 40K |
| **Image resolution** | 0.5 µm / pixel | ∼1.0 × 1.0 × 1.0 mm / voxel | ∼ 1.0 × 1.0 × 1.0 mm / voxel | NA | ∼1.0 × 1.0 × 1.0 mm / voxel | ∼0.02 mm / pixel | NA |
| **Input dimension** | 10, 000 x 2048 | 128 x 128 x 128 | 48 x 60 x 48 | 39 | 64 x 192 x 192 | 200 x 200 x 3 | 13 |

# Deploying a FL system

## Challenges




Framework

Fairness

Explainability/
Interpretability

# Deploying a FL system
## Challenges

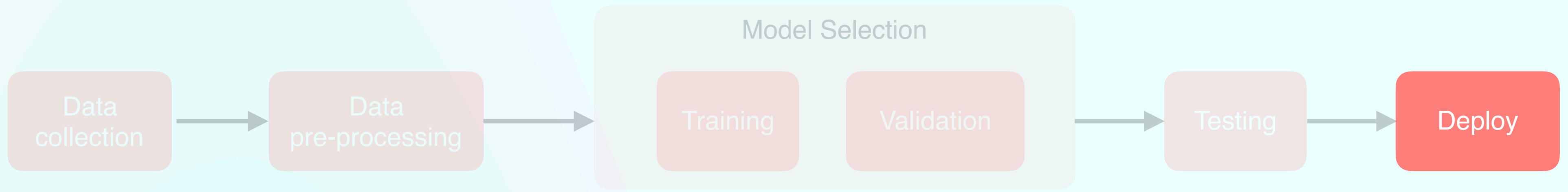
Data collection → Data pre-processing → Model Selection (Training, Validation) → Testing → Deploy



Framework

Fairness

Explainability/
Interpretability

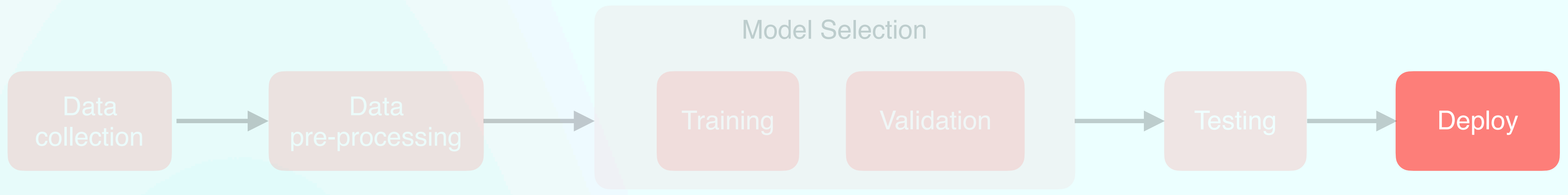



# Interpretability
## Challenge

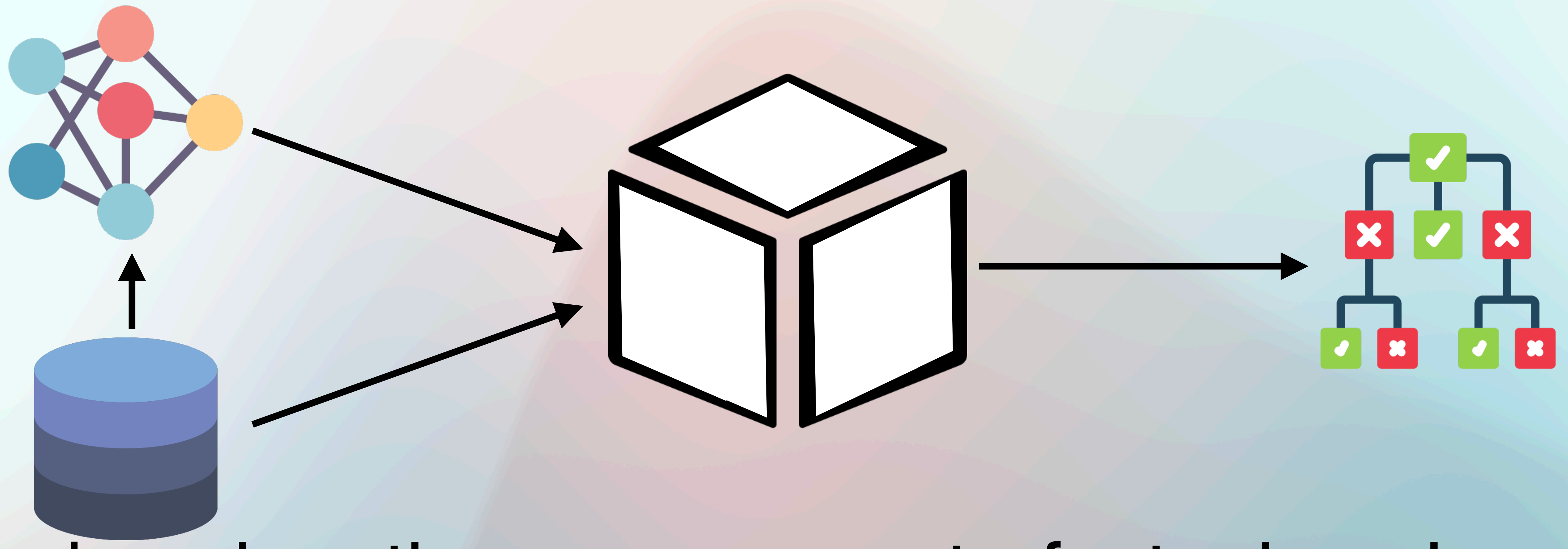

- **Neural network** models are generally **hard to interpret** (exceptions could be made for images)
- In medical applications, decisions affect the lives of human beings and a black-box machines cannot be blindly trusted!

# Interpretability

## Research directions




- Client-side surrogate models (model agnostic)



- Example-based explanations, e.g., counterfactual explanation or k-NN

- Extend the FL framework to non gradient-based methods

# Non gradient-based FL

## Federated Adaboost

Roberto Esposito, **Mirko Polato**, Marco Aldinucci. Boosting Methods for Federated Learning. SEBD 2023

Roberto Esposito, **Mirko Polato**, Marco Aldinucci. Boosting the federation: Cross-silo federated learning without gradient descent. IJCNN 2022

Gianluca Mittone, Walter Riviera, Iacopo Colonnelli, Robert Birke, Marco Aldinucci. Model-agnostic Federated Learning. Euro-Par 2023

THE LANCET


Volume 397, Issue 10270, 16–22 January 2021, Pages 199-207

Articles

Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): a modelling study of pooled datasets

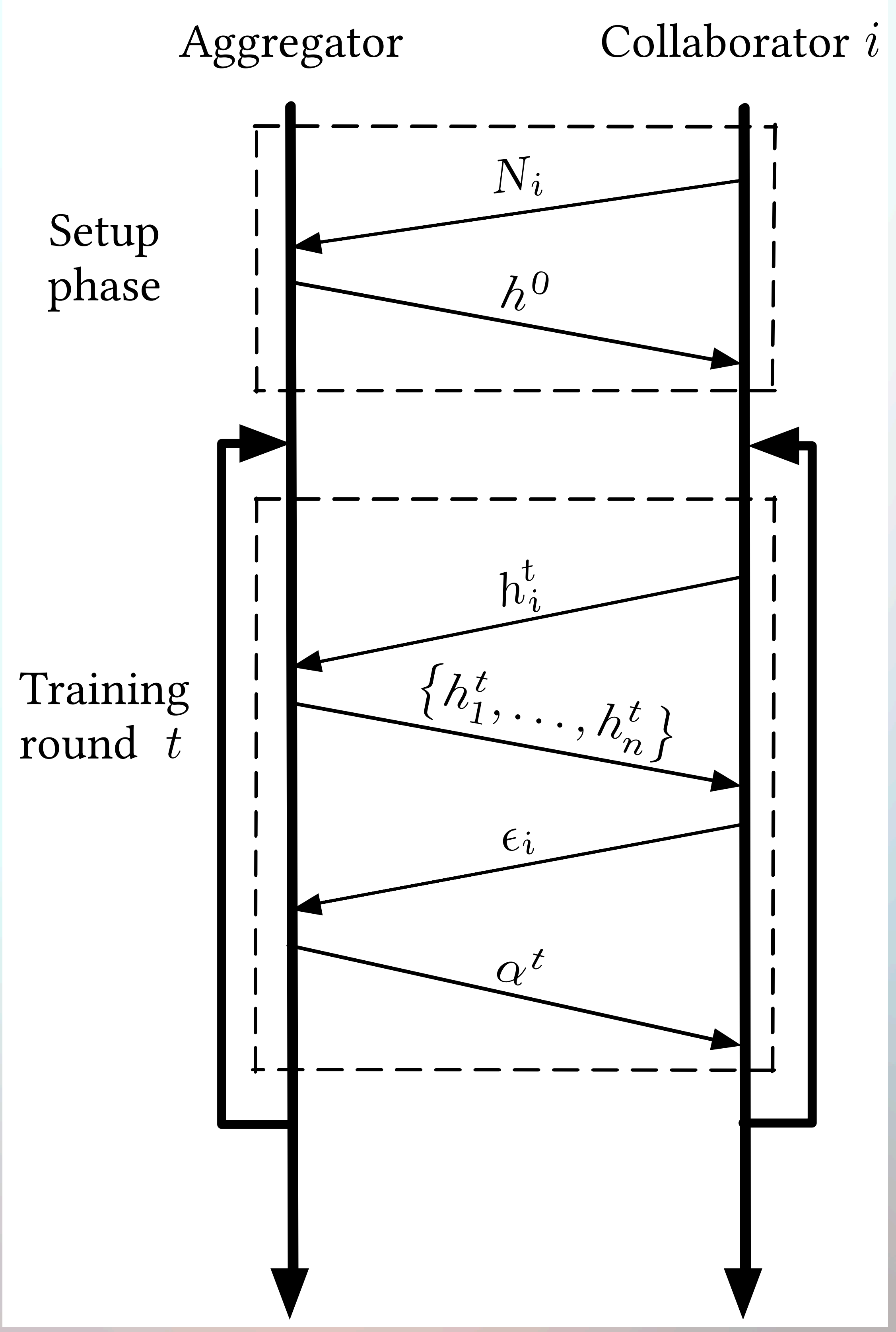
Fabrizio D'Ascenzo MD [a b], Ovidio De Filippo MD [a b], Guglielmo Gallone MD [a b], Gianluca Mittone MSc [c], Prof Marco Agostino Deriu PhD [p], Mario Iannaccone MD [i], Albert Ariza-Solé MD [f], Prof Christoph Liebetrau MD [g], Sergio Manzano-Fernández MD [h], Giorgio Quadri MD [l], Tim Kinnaird MD [e], Prof Gianluca Campo MD [o], Jose Paulo Simao Henriques MD [j], James M Hughes PhD [n], Alberto Dominguez-Rodriguez MD [m], Prof Marco Aldinucci PhD [c], Prof Umberto Morbiducci PhD [p], Prof Giuseppe Patti MD [k], Sergio Raposeiras-Roubin MD [d], Emad Abu-Assi MD [d]…Yasir Arfat

# Federated Adaboost
## Algorithms overview

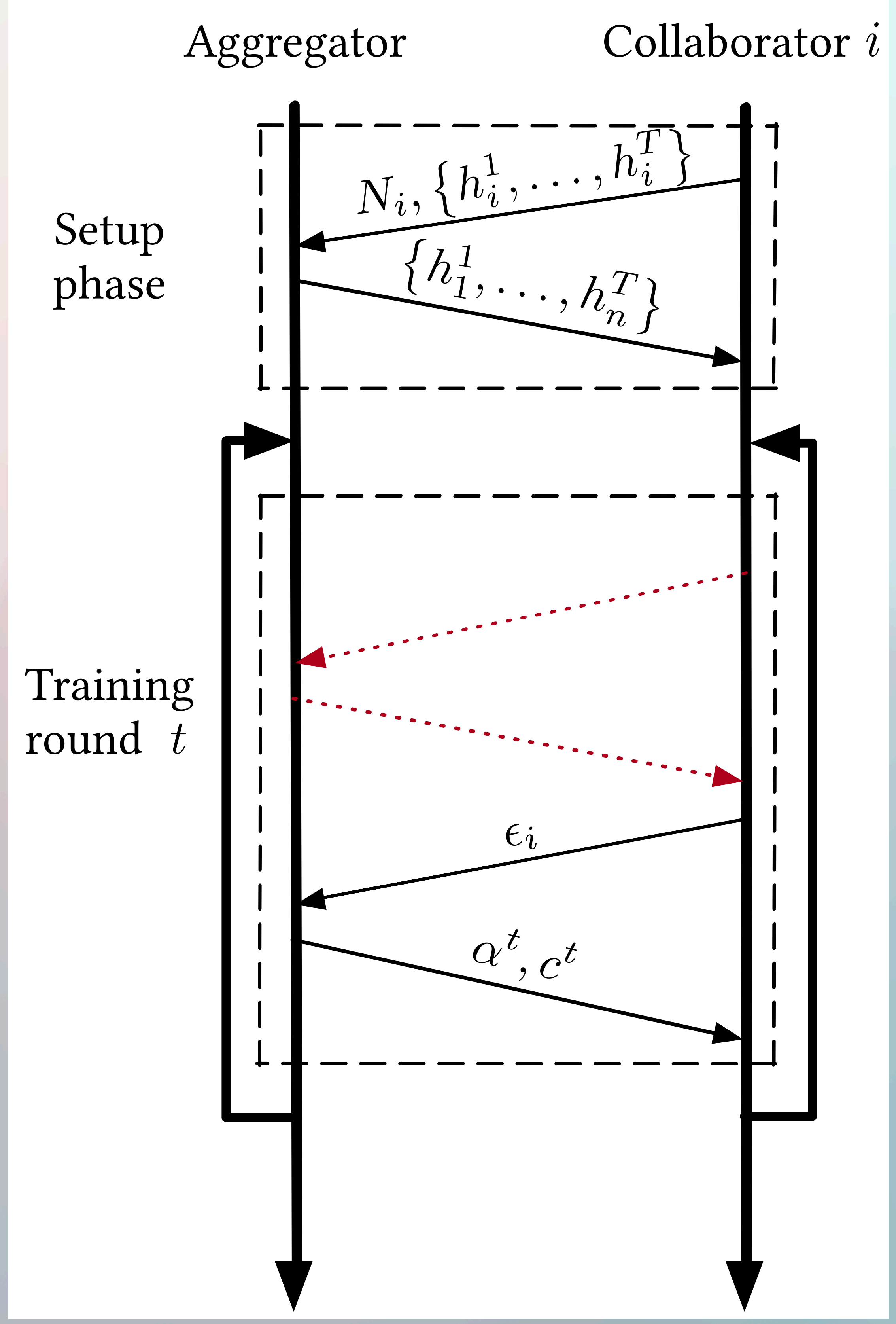
Roberto Esposito, **Mirko Polato**, Marco Aldinucci. Boosting Methods for Federated Learning. SEBD 2023

Roberto Esposito, **Mirko Polato**, Marco Aldinucci. Boosting the federation: Cross-silo federated learning without gradient descent. IJCNN 2022

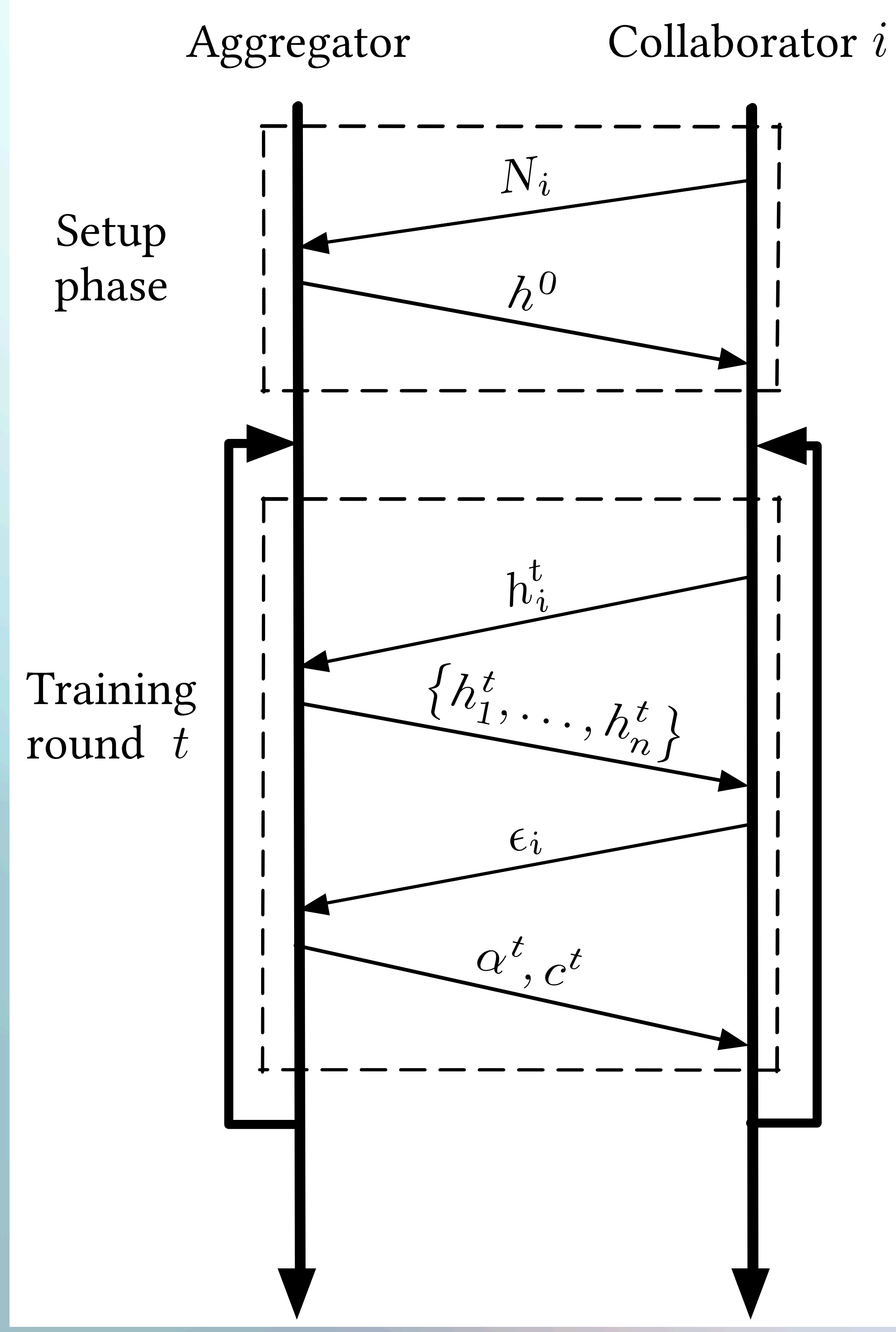
Gianluca Mittone, Walter Riviera, Iacopo Colonnelli, Robert Birke, Marco Aldinucci. Model-agnostic Federated Learning. Euro-Par 2023




**Distboost.F**




**Preweak.F**



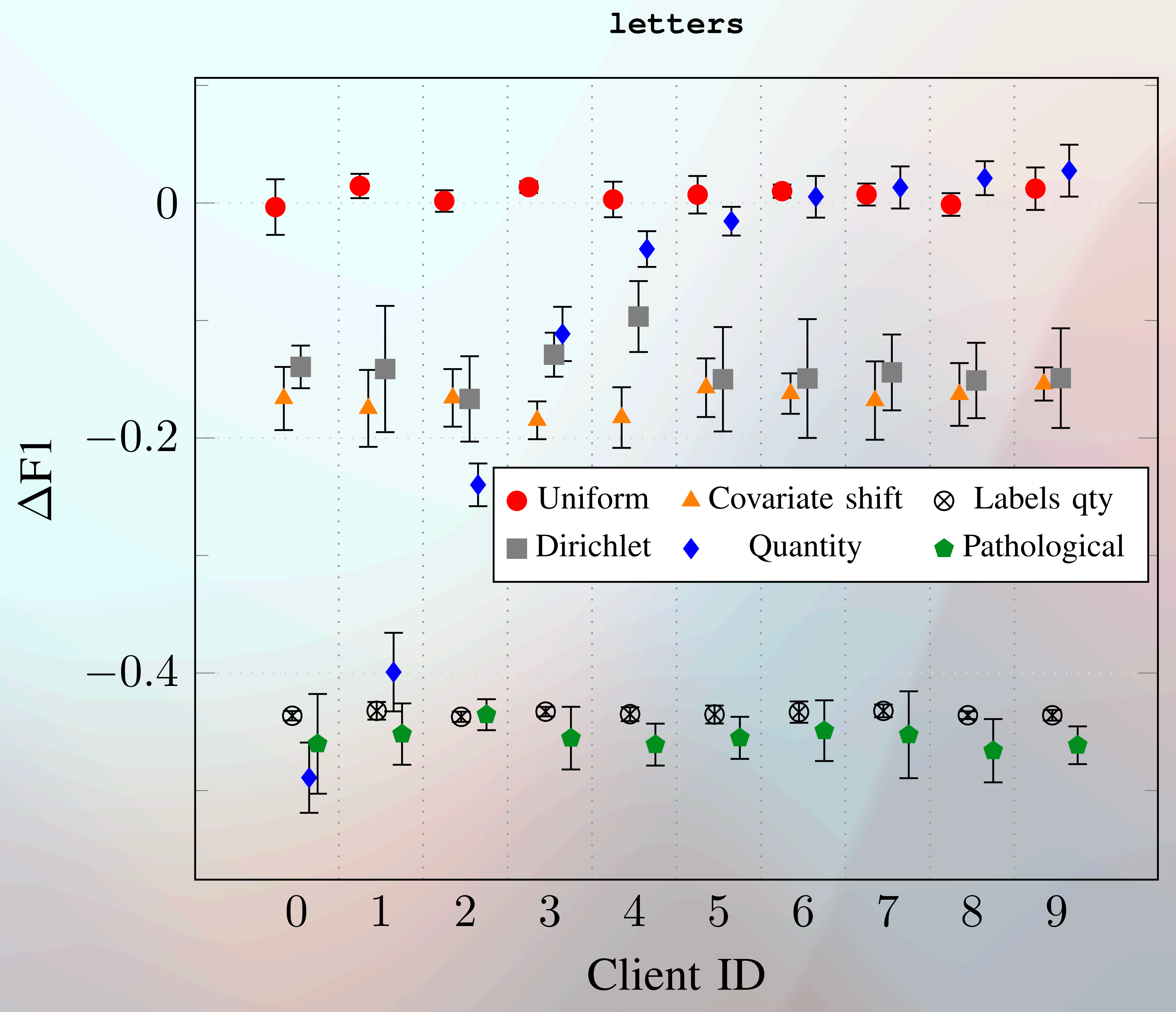

**Adaboost.F**

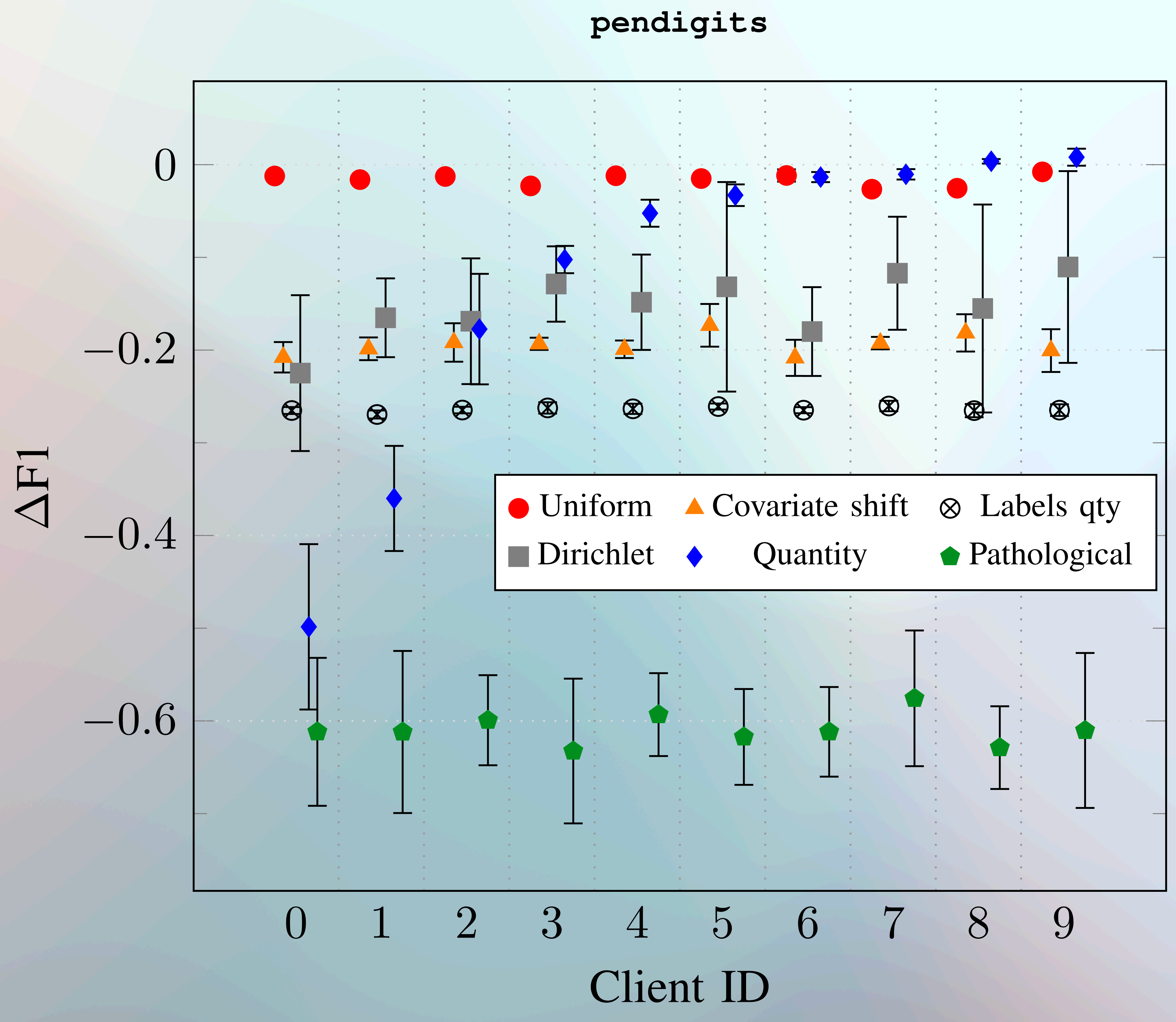
# Federated Adaboost

## Results overview

Roberto Esposito, **Mirko Polato**, Marco Aldinucci. Boosting Methods for Federated Learning. SEBD 2023

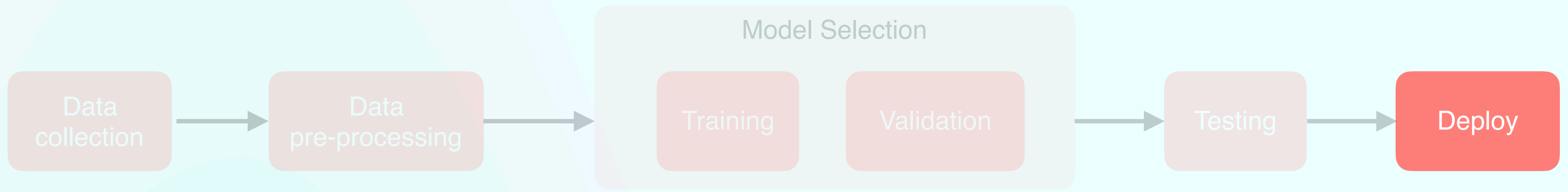
Roberto Esposito, **Mirko Polato**, Marco Aldinucci. Boosting the federation: Cross-silo federated learning without gradient descent. IJCNN 2022



# Deploying a FL system

## Challenges





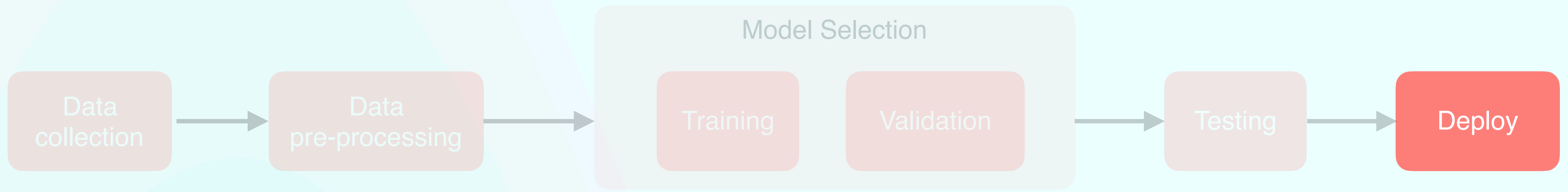

Framework

Fairness



Explainability/
Interpretability

# Fairness
## Challenges



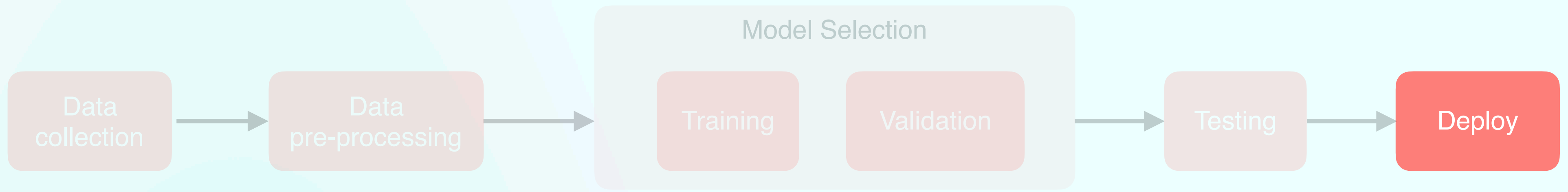

Bias

Collaboration

Performance

# Fairness

## Challenges






Bias

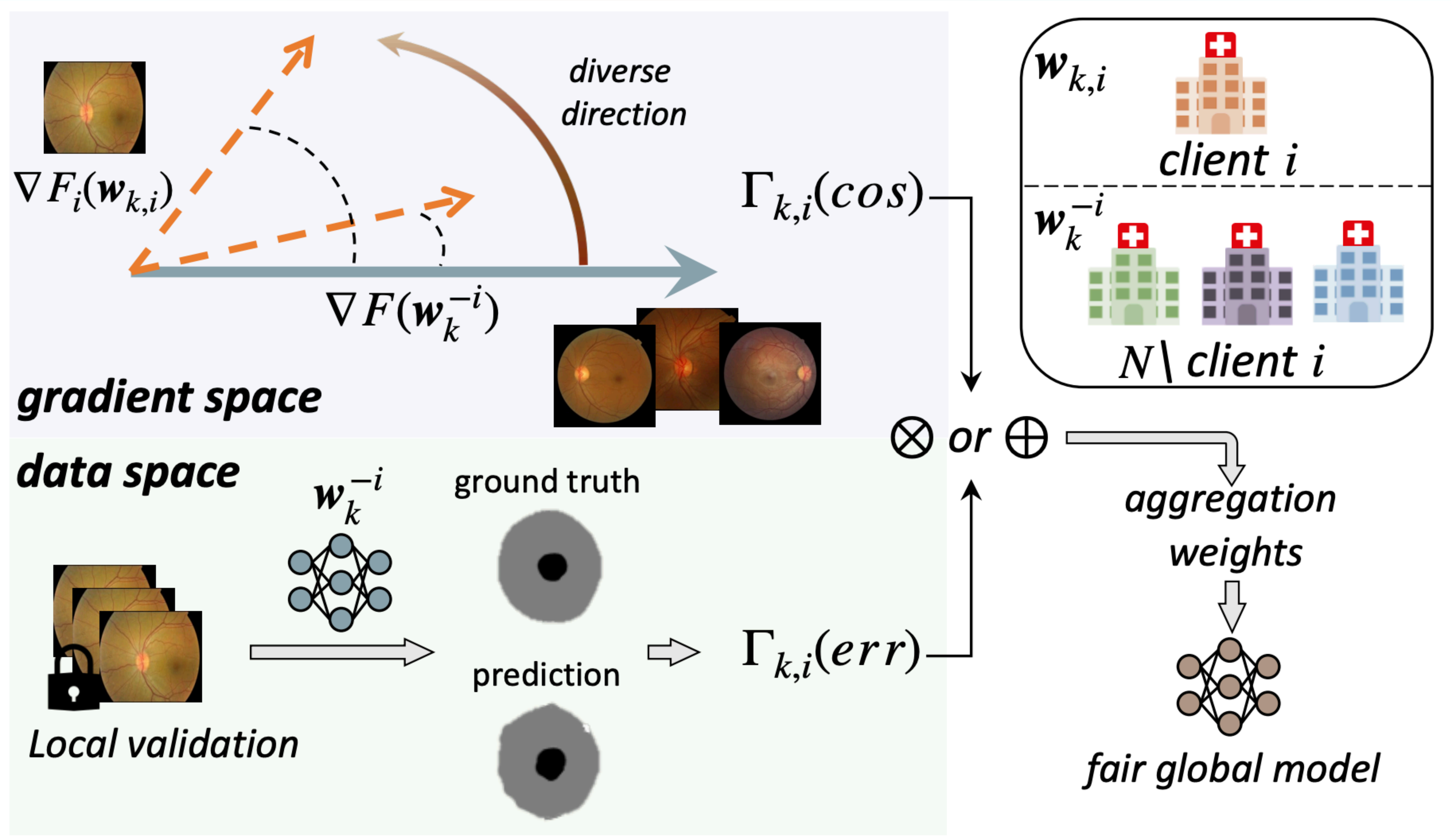
Collaboration

Performance

# Fairness

## Contribution and performance fairness

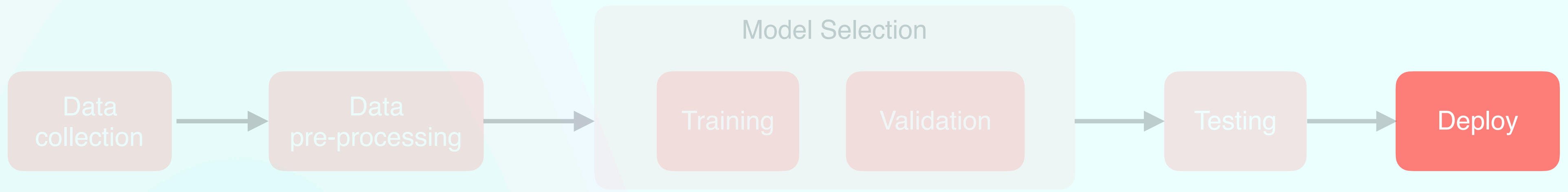
**FedCE**




Meirui Jiang, et al. Fair Federated Medical Image Segmentation via Client Contribution Estimation. CVPR 2023

# Deploying a FL system

## Challenges



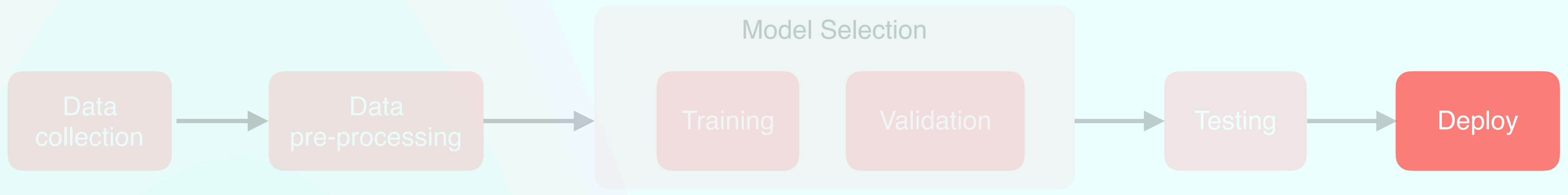

Framework

Fairness



Explainability/
Interpretability

# FL frameworks




## Many promising open source frameworks

NVIDIA FLARE

Syft

FATE

OPENFL intel

Fed-BioMed

Flower

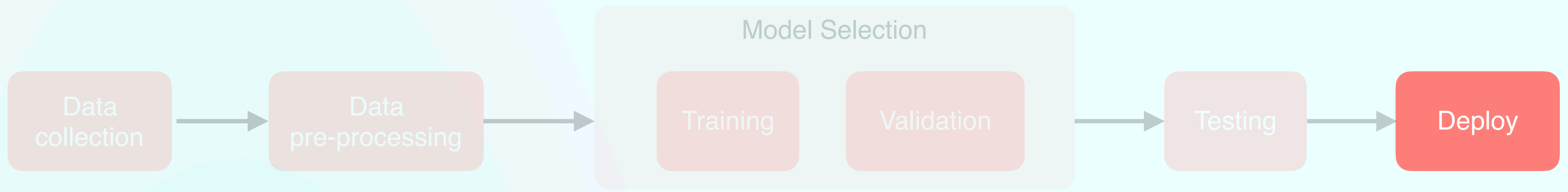
TensorFlow FEDERATED

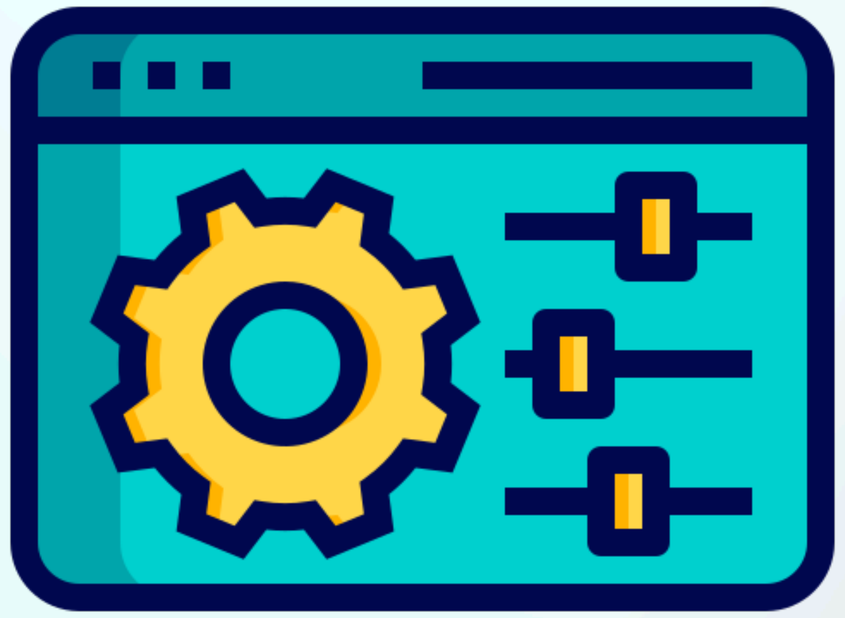
FedML

…and many others

# FL frameworks

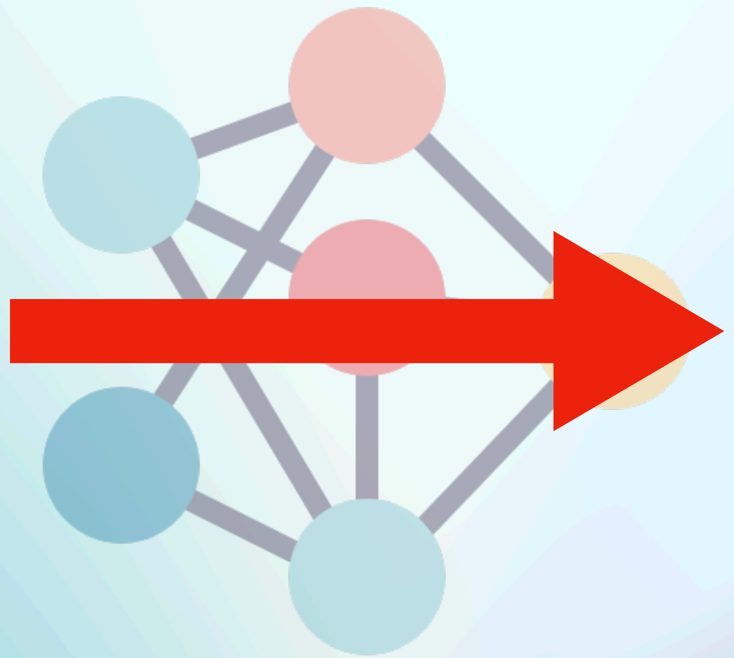
## Research directions



Configurability

Extensibility

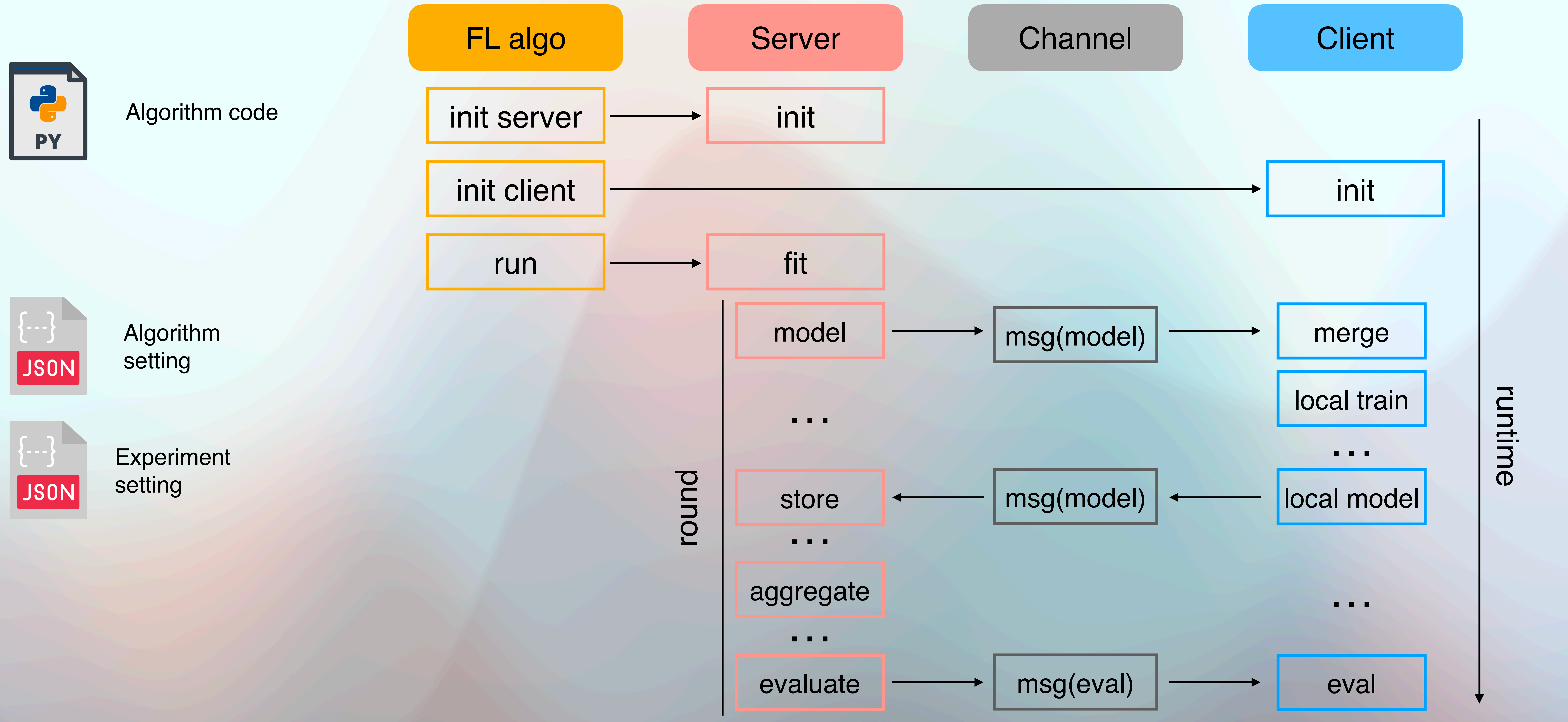


Beyond NN
(gradient descent-based)

# FL-bench (to be released soon…)

**Mirko Polato**, Roberto Esposito et al. FL-Bench
https://github.com/makgyver/fl-bench/

## Easy-to-configure & extend framework for simulated FL



# FL-bench (to be released soon…)

**Mirko Polato**, Roberto Esposito et al. FL-Bench
https://github.com/makgyver/fl-bench/

## Configuration files

### Experiment setting

```json
{
    "protocol": {
        "n_clients": 100,
        "n_rounds": 200,
        "eligible_perc": 0.1
    },
    "data": {
        "dataset": "mnist",
        "standardize": false,
        "distribution": "iid",
        "validation_split": 0.0,
        "sampling_perc": 1.0
    },
    "exp": {
        "seed": 5,
        "device": "auto",
        "checkpoint": {
            "save": false,
            "load": false,
            "path": "./checkpoints/myckp.pt"
        }
    },
    "log": {
        "logger": "local",
        "wandb_params": {
            "project": "my-proj",
            "entity": "my-entity",
            "tags": ["my-tag"]
        }
    }
}
```

### Algorithm (FedAvg) setting

```json
{
    "name": "fedavg",
    "hyperparameters": {
        "server": {

        },
        "client": {
            "batch_size": 50,
            "n_epochs": 10,
            "loss": "CrossEntropyLoss",
            "optimizer": {
                "lr": 0.1,
                "scheduler_kwargs": {
                    "step_size": 1,
                    "gamma": 1
                }
            }
        },
        "model": "MLP"
    }
}
```

# Thank you!

## An incomplete list of my colleagues…

**Mirko Polato**
Assistant Professor
University of Torino
mirko.polato@unito.it

**Roberto Esposito**
Associate Professor
University of Torino

**Marco Aldinucci**
Full Professor
University of Torino

**Samuele Fonio**
PhD Student
University of Torino

**Bruno Casella**
PhD Student
University of Torino

**Gianluca Mittone**
PhD Student
University of Torino