Lecture 13 - Contents

An overview of the main sections in this lecture.

This outline is for guidance. Navigate the slides with the left/right arrow keys.

Lecture 13:

Federated Learning for Medical LLMs

Privacy-Preserving Medical AI

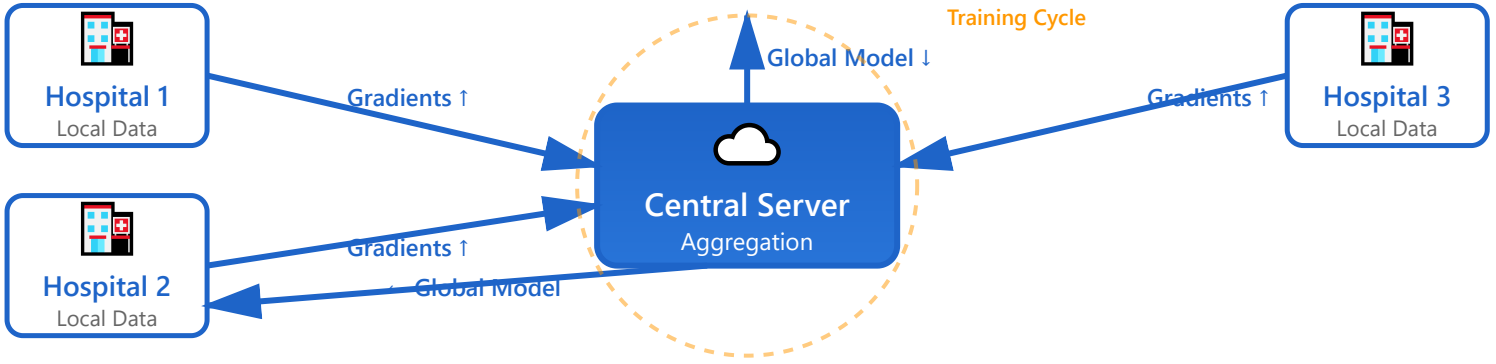
Federated Learning Overview

FL Principle

Collaborative learning without centralizing data. Models travel to data, not vice versa.

Medical Privacy

HIPAA & GDPR compliant. Patient data never leaves the hospital premises.



Data stays distributed • Models aggregate learnings • Privacy preserved

Part 1

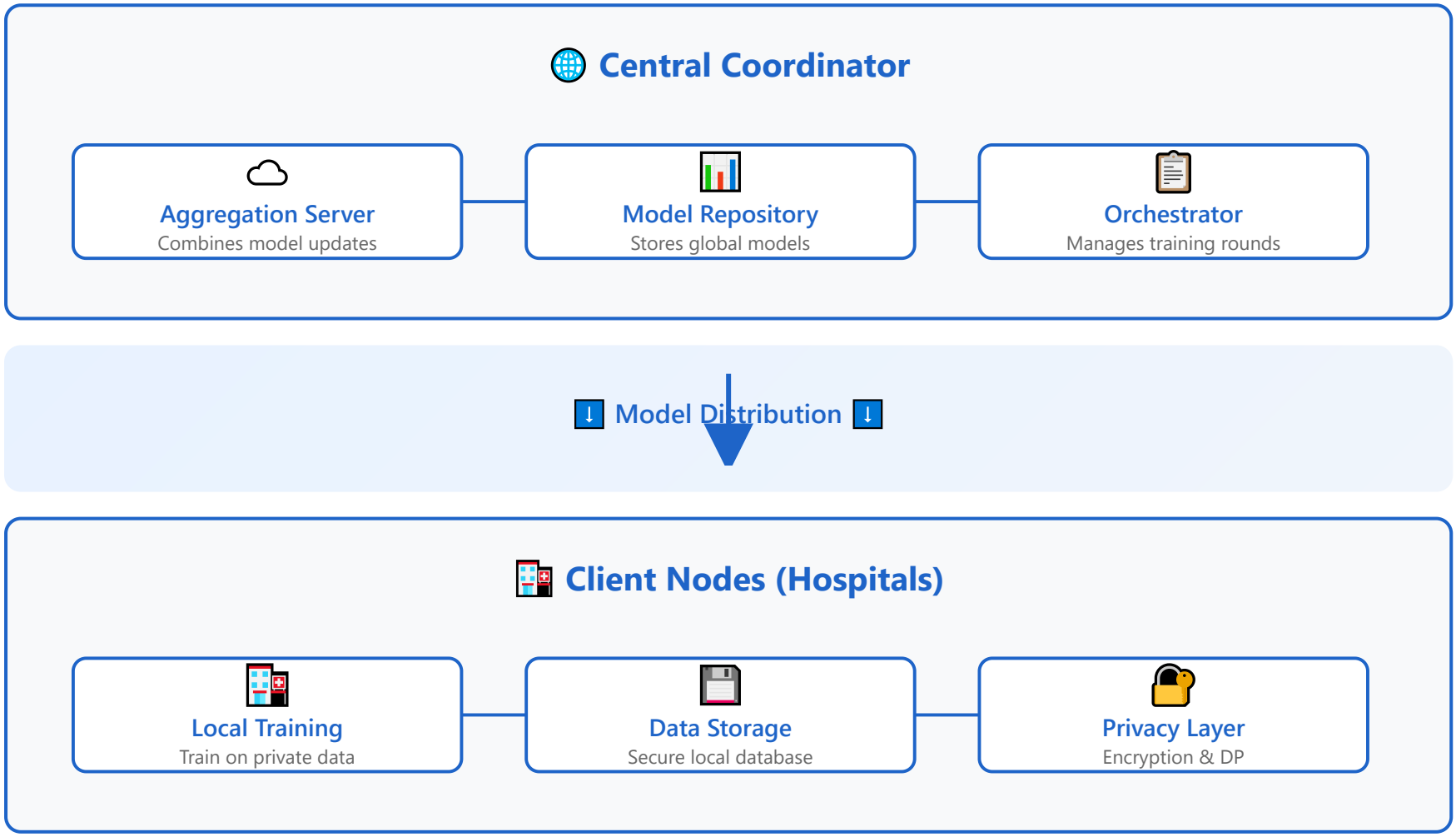
Privacy-Preserving Techniques

Encryption

Anonymization

Secure Aggregation

Distributed Architecture




Gradient Upload

Client-Server Communication

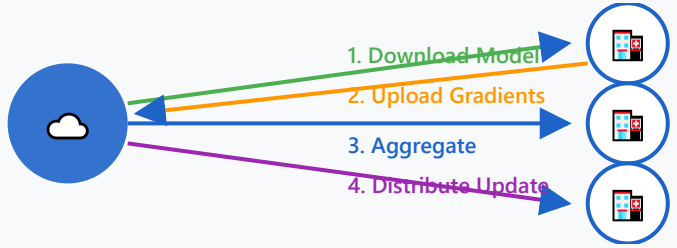
Communication Protocol

gRPC: High-performance RPC framework

REST API: HTTP-based communication

WebSocket: Real-time bidirectional

Message Flow



Security Measures

TLS 1.3

Mutual Auth

Rate Limiting

Audit Logging

Verification

Aggregation Algorithms

Weighted Average

$w_{global} = \Sigma(n_i/N \times w_i)$

w_1 w_2 w_3

Weight by dataset size

FedAvg

Most popular algorithm

★ ★ ★ ★ ★

Simple & effective

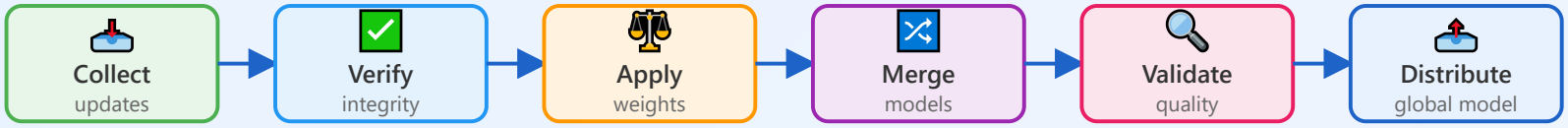
FedProx

Adds proximal term

\+ μ term

Handles heterogeneity

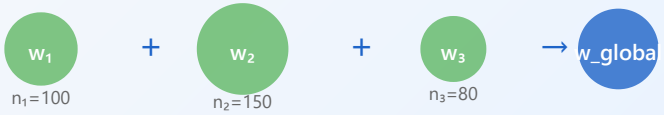
Aggregation Process



FedAvg & FedProx

FedAvg

$$w = \Sigma(n_k/n \times w_k)$$



- Simple weighted average
- Fast convergence
- Works well with IID data

FedProx

$$\min F(w) + \mu/2||w-w_t||^2$$

Standard
Loss F(w)

+

Proximal Term
$\mu/2||w-w_t||^2$

- Adds proximal term
- Handles heterogeneity
- More robust

Differential Privacy

Original Data + Calibrated Noise = Protected Data

Mechanisms

Gaussian Mechanism

Laplace Mechanism

Privacy Budget (ε)

ε remaining

Parameters

ε
Privacy Loss
Lower = More Private

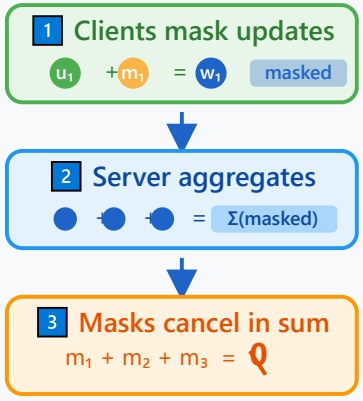
δ
Failure Probability
Typically 10^{-5}

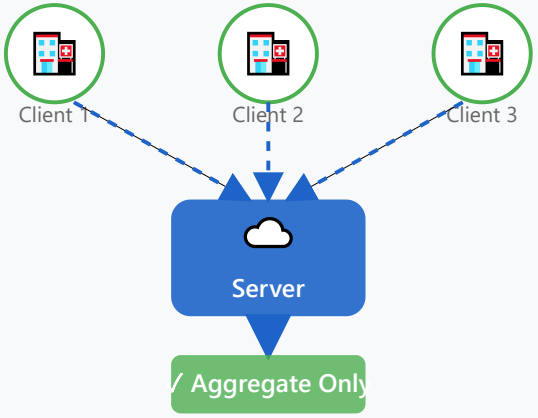
Typical Values
ε = 1-10
δ = 10^{-5}

Secure Aggregation

Server learns only aggregate, not individual updates

Process





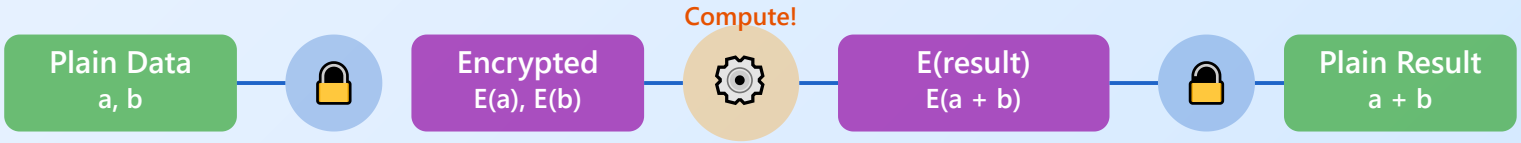
Individual updates remain secret

Shamir Secret Sharing
Split secrets into shares

Homomorphic Encryption
Compute on encrypted data

Secure Multi-Party Computation
Joint computation without revealing

Homomorphic Encryption



How It Works

Mathematical Property:

$E(a) \oplus E(b) = E(a \oplus b)$

Operation on encrypted data = Encryption of operated data

Types

Partial HE
Only addition OR multiplication

Somewhat HE
Limited number of operations

Fully HE ★
Arbitrary computations

Security

100-1000x slowdown

Performance

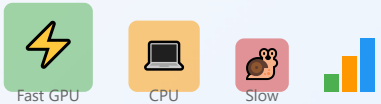
Part 2

Medical Data Challenges

Data Heterogeneity

Statistical

Different data distributions across hospitals



System

Varying compute resources and network

Model

Different architectures and objectives

Non-IID Medical Data

⚠ **Different patient demographics**
Age, gender, ethnicity vary by location

⚠ **Varying disease prevalence**
Regional differences in disease patterns

⚠ **Hospital-specific protocols**
Different treatment guidelines and standards

⚠ **Equipment differences**
Imaging devices, sensors vary in quality

⚠ **Regional health patterns**
Climate, lifestyle, socioeconomic factors

Impact: Model bias • Slower convergence • Reduced accuracy

Client Drift Handling

FedProx

Proximal term limits drift

SCAFFOLD

Variance reduction

FedDyn

Dynamic regularization

Adaptive LR

Adjust learning rates

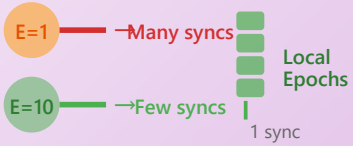
Communication Efficiency

Gradient Compression

Sparsification, quantization

10-100x reduction

Local Updates



More local epochs

Linear in E reduction

Model Compression

Pruning, distillation

2-10x reduction

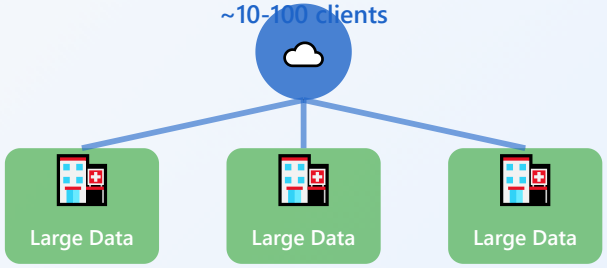
Model Personalization

- ✓ Fine-tune last layers locally
- ✓ Meta-learning (MAML)
- ✓ Mixture of global & local models
- ✓ Clustered federated learning

Global knowledge ⚖ Local adaptation

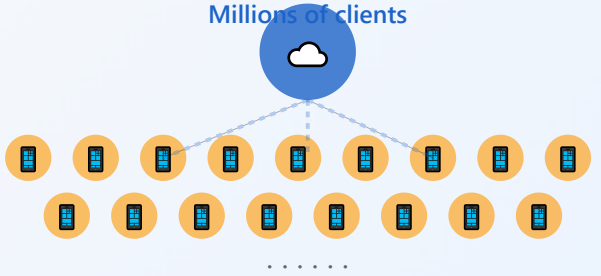
Cross-Silo vs Cross-Device

Cross-Silo (Hospital FL)



- Few clients (~10-100)
- Large datasets
- Reliable connectivity
- Powerful compute

Cross-Device (Mobile FL)



- Many clients (millions)
- Small local data
- Unreliable connectivity
- Limited resources

Part 3

Multi-Hospital Collaborations

Multi-Institutional Studies

🌍 Broader Data Diversity

Access to diverse patient populations across regions and demographics

📊 Larger Sample Size

Combined datasets enable robust statistical analysis

🔒 Privacy Preserved

Collaboration without sharing raw patient data

🏆 Better Generalization

Models trained on heterogeneous data perform better

Research Outcomes

✓ Faster rare disease diagnosis

✓ Improved treatment predictions

✓ Reduced model bias

✓ Enhanced clinical decision support

Regulatory Compliance

EU GDPR

European Union General Data Protection Regulation

- Right to explanation
- Data minimization
- Privacy by design
- Cross-border transfer rules

US HIPAA

Health Insurance Portability and Accountability Act

- PHI protection
- Security safeguards
- Breach notification
- Business associate agreements

✓ Compliance Checklist

☑ Encrypted communication channels

☑ Audit logs and monitoring

☑ Data anonymization techniques

☑ Consent management

☑ Regular security assessments

Data Governance

🏛 Governance Structure

Clear roles and responsibilities

📜 Policy Framework

Data usage policies and procedures

🔑 Access Control

Role-based permissions

📊 Audit Trail

Complete logging of data access

Quality Control

✓ Data Validation

Schema validation and quality checks

🧪 Model Testing

Comprehensive evaluation metrics

📈 Performance Monitoring

Continuous accuracy tracking

🔍 Bias Detection

Fairness and equity assessments

Incentive Mechanisms

💰 Financial Rewards

Payment for data contribution

🏆 Reputation System

Recognition and credibility

📊 Data Credit

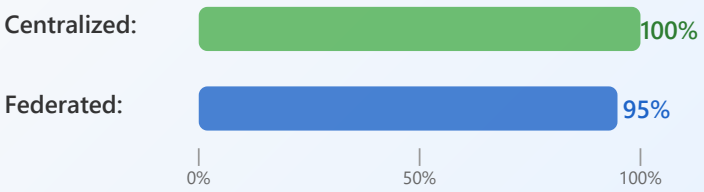
Fair attribution of contributions

🎯 Performance Bonus

Rewards for quality data

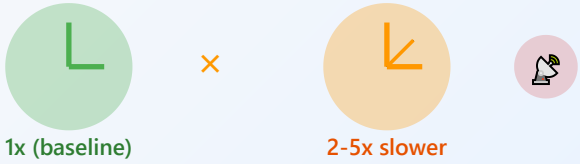
Performance Benchmarks

Centralized vs FL



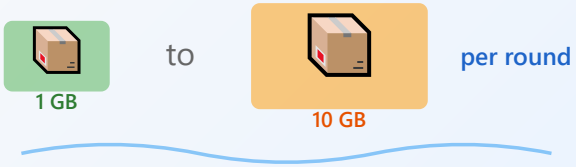
90-95% accuracy retention in FL

Training Time



2-5x slower due to communication

Bandwidth Usage



1-10 GB per round

Convergence

10-50 rounds to converge

Security Auditing

🔍 Vulnerability Analysis

Regular penetration testing

📝 Compliance Audits

GDPR/HIPAA checks

🛡️ Threat Modeling

Identify attack vectors

📊 Security Metrics

Track security posture

Scalability Analysis

📈 Node Scaling

Linear scaling up to 100s of nodes

⚡ Bottlenecks

Network & aggregation server

🔄 Load Balancing

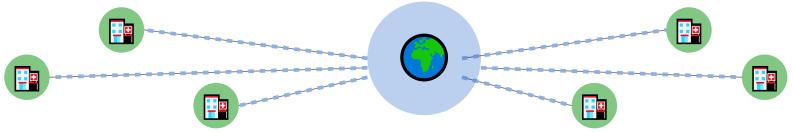
Distribute training rounds

💾 Resource Management

Optimize memory & compute

Case Study: COVID-19 FL Consortium

Global Collaboration for Pandemic Response



20+
Countries

100+
Hospitals

1M+
Patient Records

Key Achievements

- Developed COVID-19 severity prediction models
- Achieved 85%+ accuracy across diverse populations

- Rapid deployment during critical pandemic phases

- Full GDPR/HIPAA compliance maintained

- Cross-border collaboration without data transfer

Hands-On: Federated Setup

Flower

Friendly FL framework with simple API

```
pip install flwr
flwr.client.start()
```

PySyft

Privacy-preserving ML library

```
pip install syft
sy.VirtualMachine()
```

Setup Steps

1. Install framework (Flower/PySyft)
2. Define model architecture
3. Configure server & clients
4. Implement training loop
5. Run federated training
6. Evaluate global model

Future Directions

Vertical FL

Different features from different institutions

Genomic FL

Privacy-preserving genetic research

LLM Fine-tuning

Federated adaptation of medical LLMs

Split Learning

Hybrid FL + split computation

Research Opportunities

Communication efficiency • Adaptive aggregation • Cross-silo heterogeneity • Byzantine robustness • Incentive mechanisms • Real-world deployments

Thank You! 

Key Takeaways

✓ FL enables privacy-preserving multi-hospital collaboration

✓ Differential privacy & secure aggregation protect data

✓ Heterogeneity requires specialized algorithms (FedProx, SCAFFOLD)

✓ GDPR/HIPAA compliance is achievable with proper design

✓ Real-world deployments show promising results

📚 Resources: Flower Framework • PySyft • NVIDIA FLARE

📎 Papers: FedAvg • FedProx • Secure Aggregation