DIPARTIMENTO di INGEGNERIA

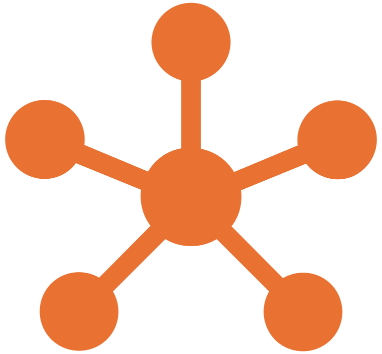
LA RICERCA DENTRO UNIME

Machine learning for environmental cyber physical systems management

Candidate: Ilenia Ficili

Tutor: Prof. Antonio Puliafito

Context of the study



Internet of Things

Machine Learning

Environmental Monitoring

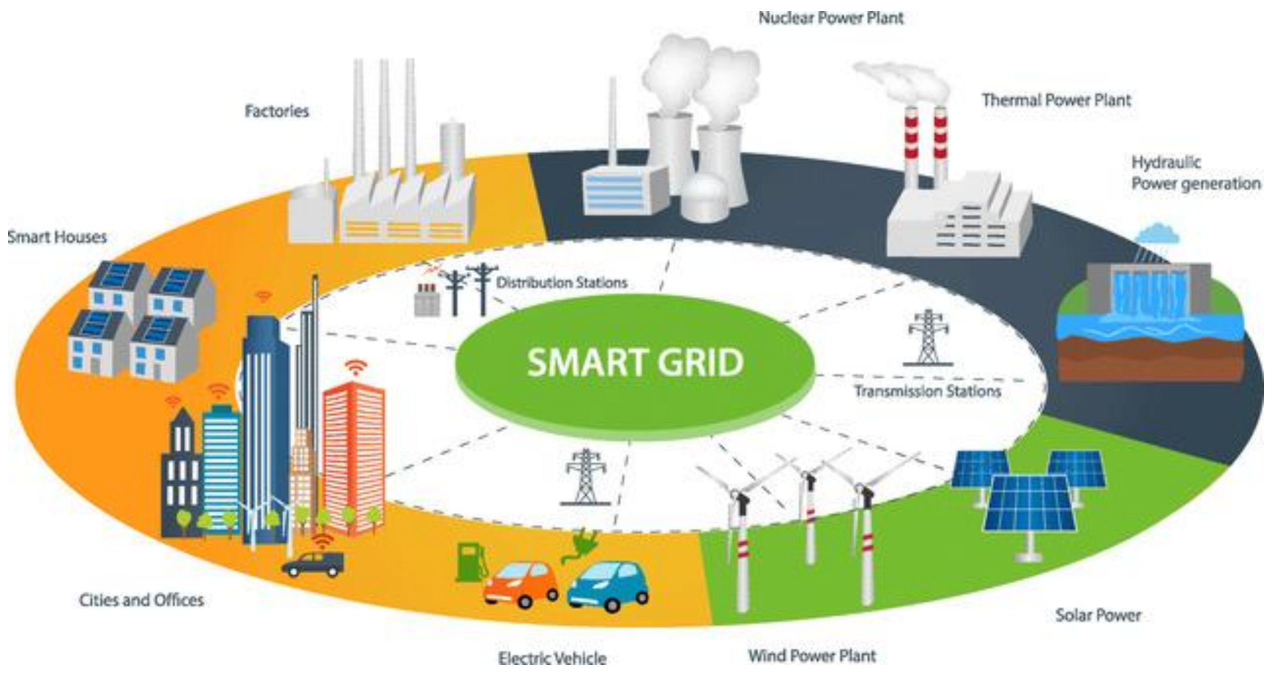
Key Components and Examples of CPS

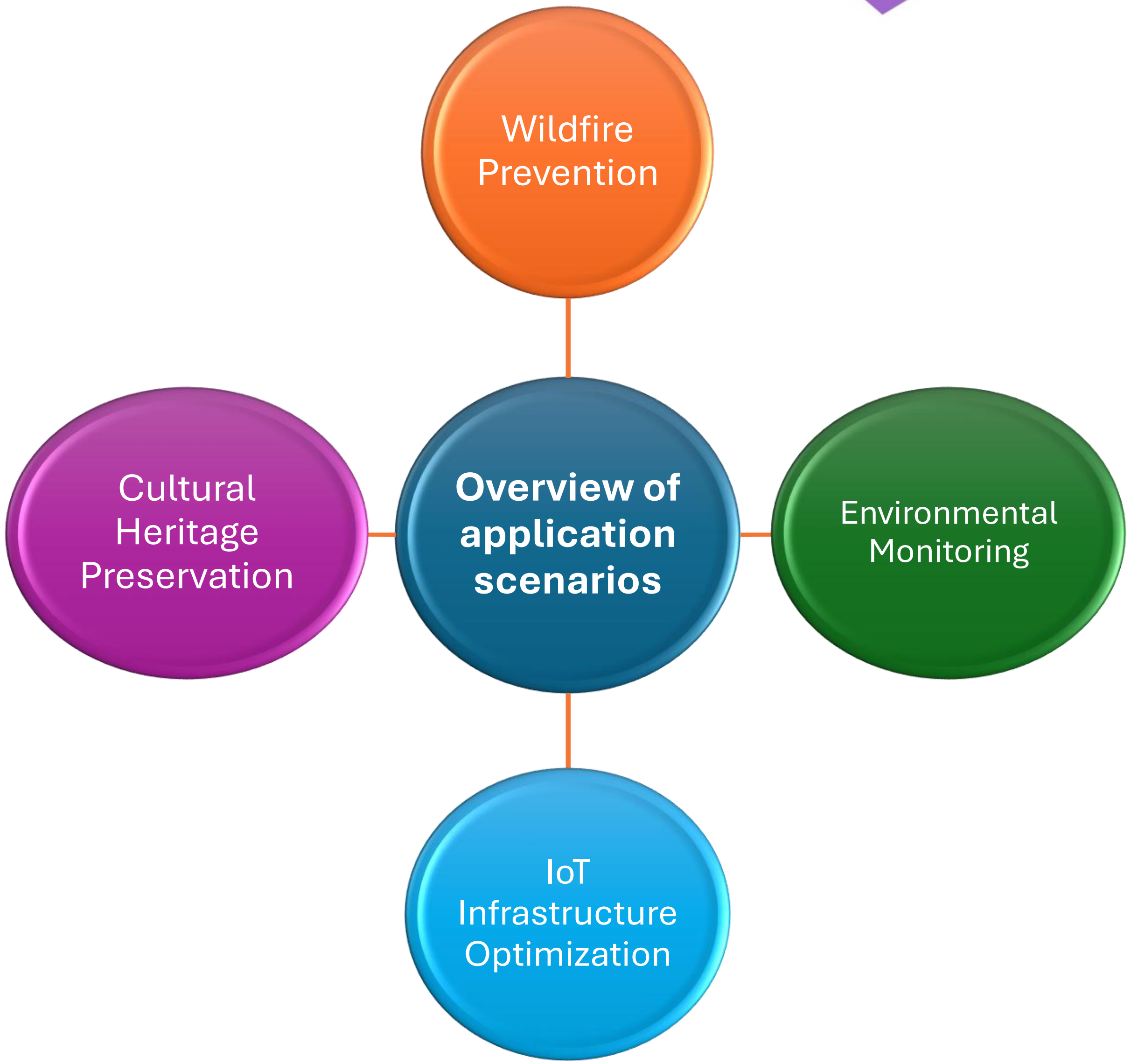
Main components of CPS:

- **Sensors** – Collect data from the environment.
- **Computing Units** – Process data and make decisions.
- **Actuators** – Execute actions based on processed information.

Examples:

- *Autonomous Vehicles*: Sensors detect surroundings, AI makes driving decisions, actuators control movement.
- *Smart Grids*: Optimize energy distribution based on real-time demand.
- *Early Warning Systems*: Detect natural disasters like floods and earthquakes.




Wildfire Prevention
Cultural Heritage Preservation
Overview of application scenarios
Environmental Monitoring
IoT Infrastructure Optimization

Core Methodologies and Technologies

- *Convolutional Neural Network*
- *Incremental Learning*
- *Federated Learning*

INCREMENTAL LEARNING

Incremental Learning (IL) allows AI models to continuously update with new data without retraining from scratch.

Why It's Important:

- Avoids **catastrophic forgetting** – prevents loss of previously learned knowledge.
- Adapts to **changing environments** – models evolve as new patterns emerge.
- Reduces **computational costs** – updates only when necessary.

How It Works

1. **Initial Training** – The model learns from an initial dataset.
2. **New Data Integration** – As new samples arrive, the model updates selectively.
3. **Knowledge Retention** – Mechanisms like memory replay or regularization ensure past knowledge is not lost.

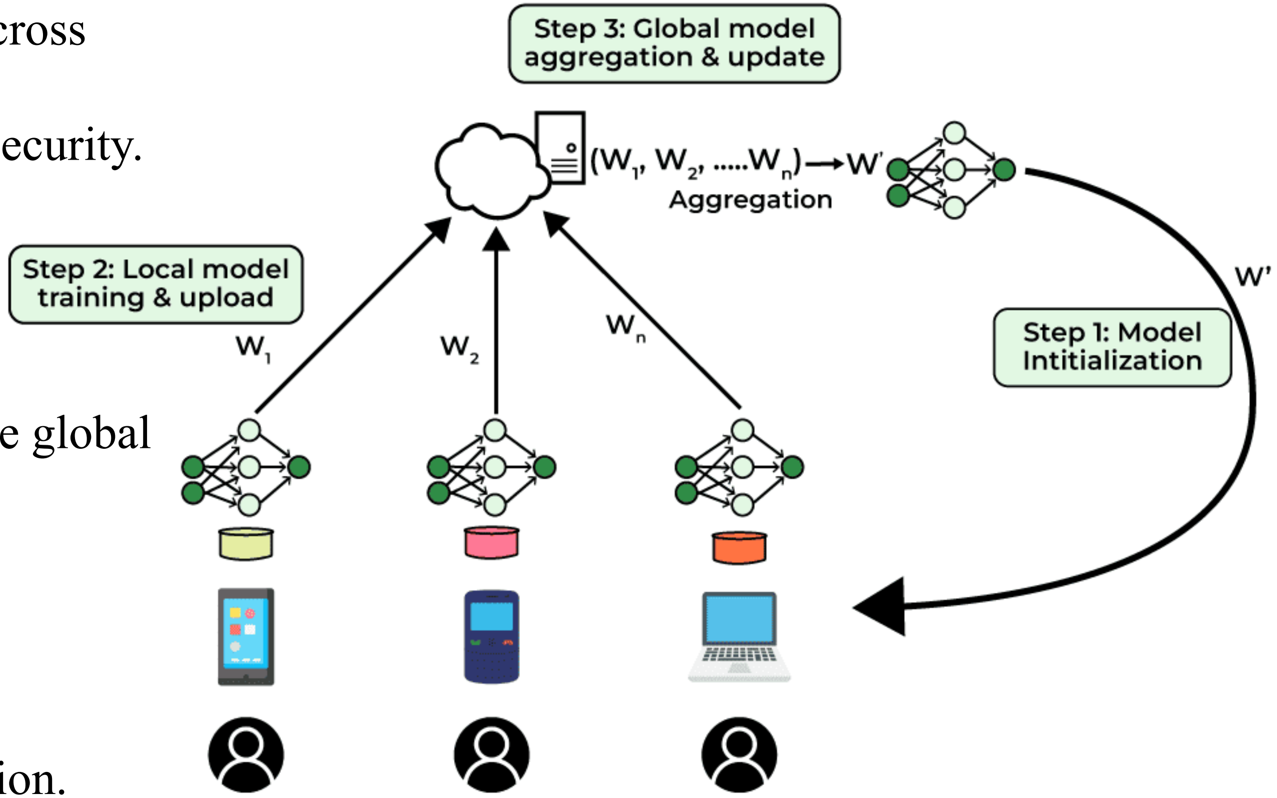
Federated Learning

What is Federated Learning?
A machine learning paradigm where models are trained across decentralized devices or servers.
Data never leaves the local device, ensuring privacy and security.

<u>How it Works:</u>

- Local devices train a model on their data.
- Only model updates (weights) are shared, not raw data.
- A central server aggregates these updates to improve the global model.

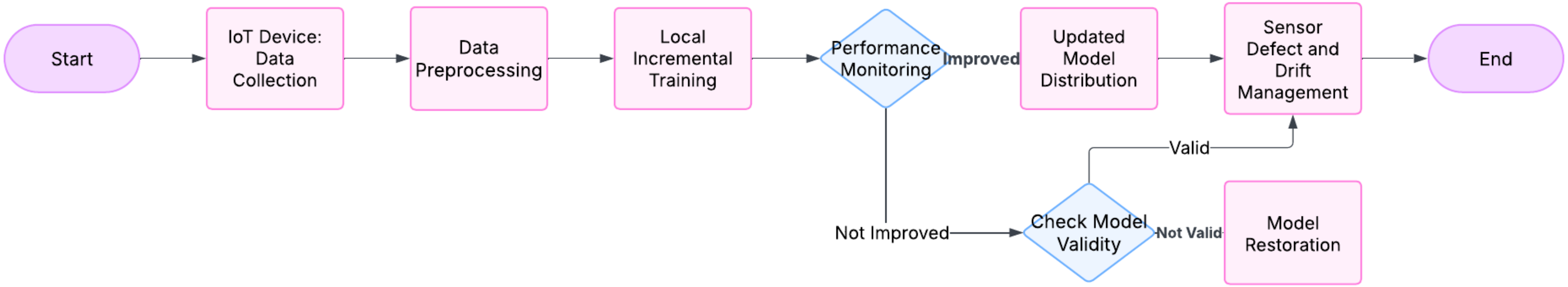
<u>Key Benefits:</u>
Data Privacy: Sensitive data stays on the device.
Efficiency: Reduces data transfer, enhancing scalability.
Security: Protects against data breaches during transmission.



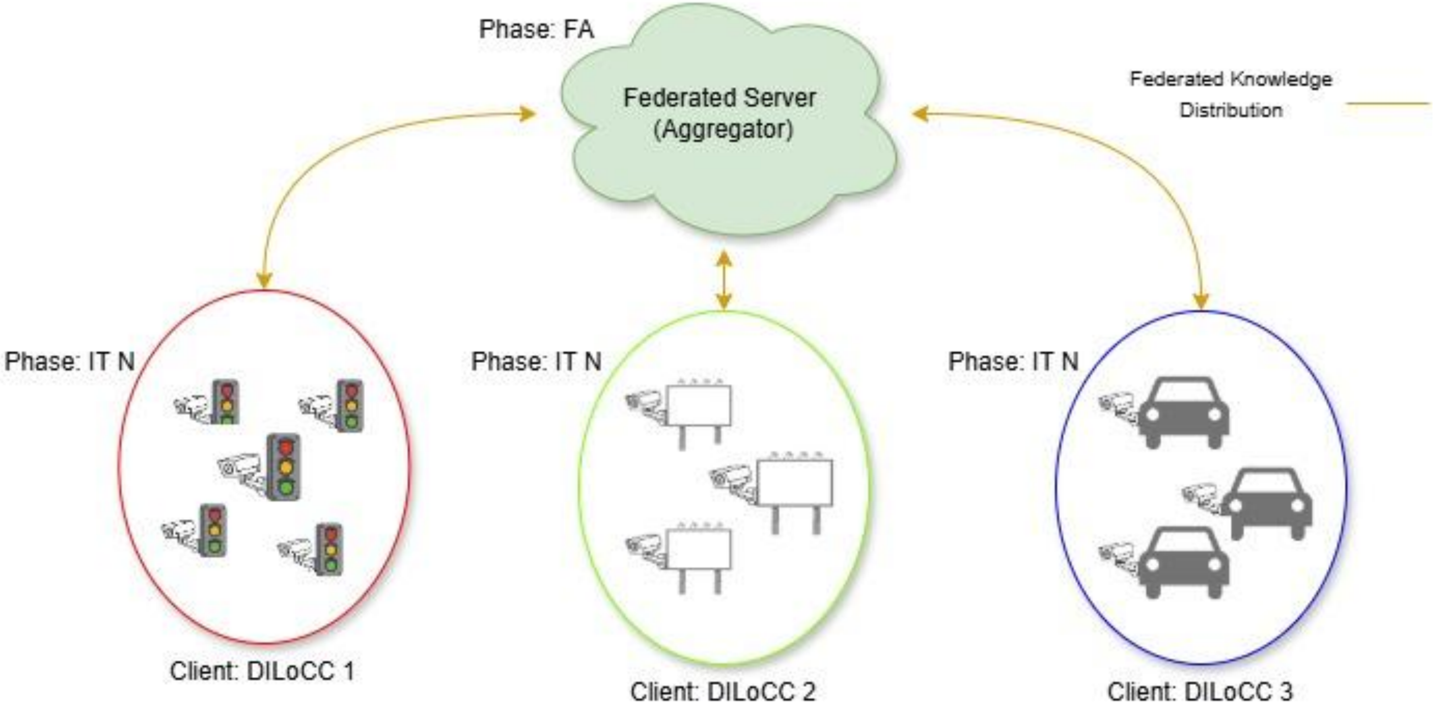
DILoCC – Distributed Incremental Learning

What is DILoCC?
Distributed Incremental Learning on Computing Continuum.
Designed for **dynamic AI model updates** on edge devices.

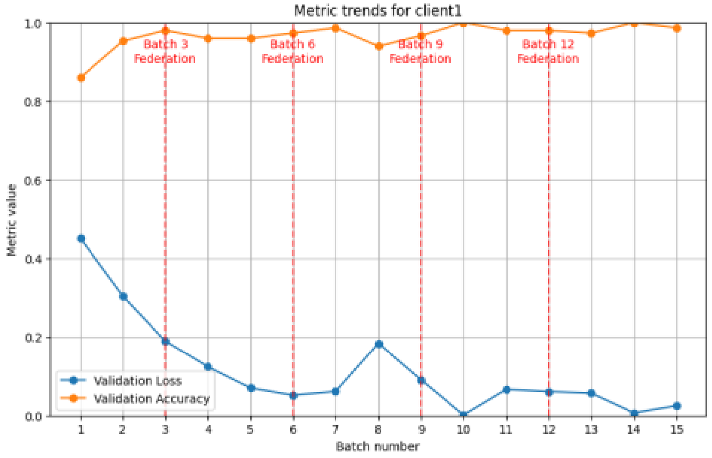
Key advantages:
Reduces **latency** and optimizes **resource consumption**.
Uses **cloud + edge computing** for scalable AI deployment.
Prevents **catastrophic forgetting**, ensuring continuous learning.

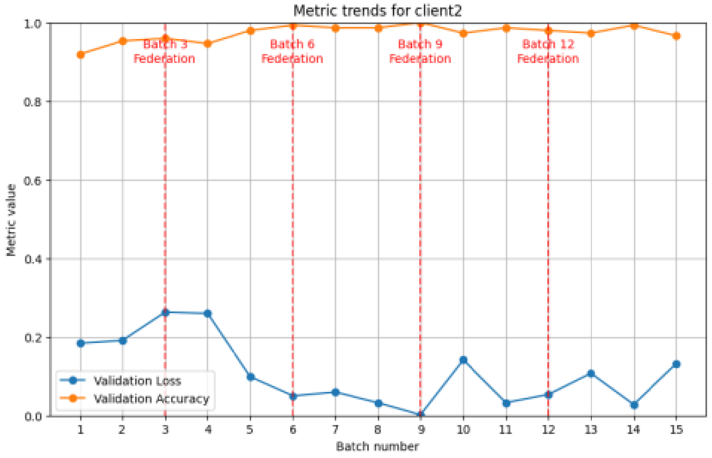


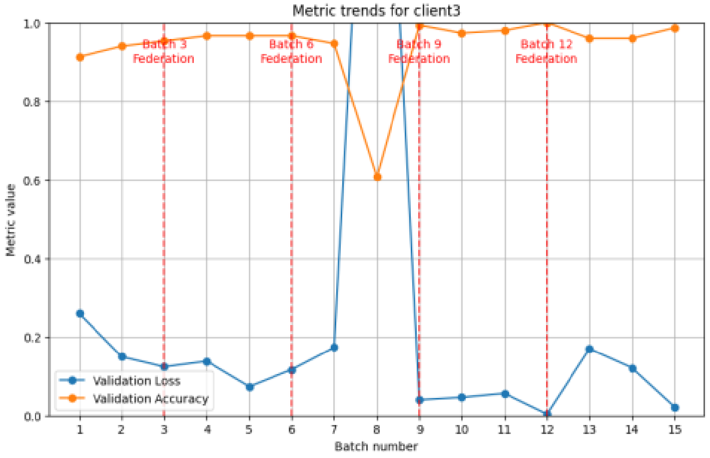
Traffic Management: a hybrid incremental - federated learning approach



Traffic Management: a hybrid incremental - federated learning approach







Metric	Base Model	Client 1	Client 2	Client 3
Accuracy (%)	79.33	97.30	98.00	98.00
Loss	0.6493	0.1591	0.2641	0.0358

HERALD: A FEDERATED-INCREMENTAL APPROACH

- The Covid-19 pandemic exposed the lack of coordination in healthcare systems.
- Infection diagnosis requires efficient tools that minimize direct physician intervention.
- Data privacy regulations hinder centralized training of ML models.
- HERALD applied to **chest X-ray images** (COVID-19 and healthy cases).

Objective:

HERALD integrates **Incremental Learning** and **Federated Learning** to:

- Adapt models to virus mutations over time.
- Enable **privacy-preserving** knowledge sharing across hospitals.
- Mitigate the **Catastrophic Forgetting** issue.

G. Tricomi, G. Cicceri, I. Ficili, S. Vitabile, G. Merlino, and A. Puliafito, "HERALD: a Hybrid distributEd leaRning incrementAL & feDerated solution for knowledge distillation in COVID-19 classification," *Future Generation Computer Systems*, 2025. (submitted)

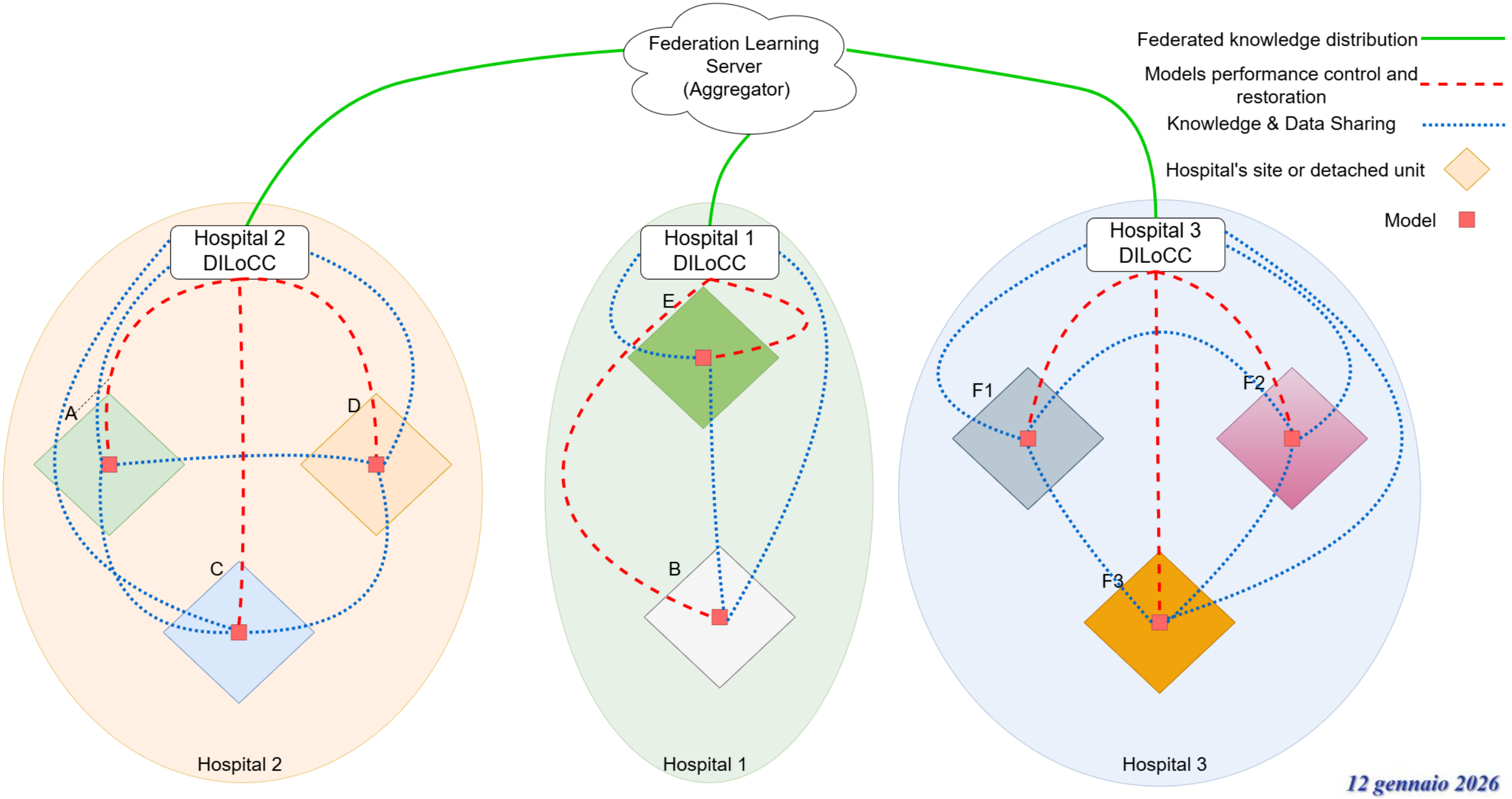
HERALD: A FEDERATED-INCREMENTAL APPROACH

Hybrid Approach:

- **Incremental Learning (IL):** continuous model updates without forgetting previous knowledge.
- **Federated Learning (FL):** decentralized collaboration without sharing raw data.

Results

- **Improved** accuracy, recall, and loss compared to traditional models.
- **Balanced** local adaptation and global knowledge integration.



Using the Knowledge Distillation techinque

Knowledge Distillation is a technique used to transfer knowledge from a large, complex model (Teacher) to a smaller and more efficient model (Student).

- The **Teacher model** is usually accurate but computationally expensive
- The **Student model** is smaller and faster, suitable for deployment.
  The Student learns not only from ground-truth labels, but also from the soft predictions of the Teacher

Goal: obtain a lightweight model with performance close to the Teacher

How Knowledge Distillation Works

The Teacher produces a probability distribution over classes
A **softmax with temperature (T)** is used to smooth the output probabilities
The Student is trained using a combination of:

- **Standard classification loss** (e.g., Cross-Entropy)
- **Distillation loss** (e.g., KL Divergence between Teacher and Student outputs)

This allows the Student to better generalize and mimic the Teacher's behavior

$$\mathcal{L} = \alpha\mathcal{L}_{CE} + (1 - \alpha)\mathcal{L}_{KD}$$

Benefits:
Reduced model size and inference cost
Comparable accuracy
Well-suited for ***federated and edge learning scenarios***

Thanks for your attention.

Ilenia Ficili
ilficili@unime.it