## 1. General

The rise of IoT sensors has enabled applications like smart cities, environmental monitoring, and precision medicine by generating massive data critical for addressing challenges like climate change. However, traditional centralized computing struggles with the sheer data volume, privacy concerns, and latency, necessitating efficient Edge-AI systems for resource-constrained devices. Federated learning (FL) addresses data privacy by enabling decentralized model training, but its effectiveness diminishes with heterogeneous, non-IID data distributions and incurs high communication costs in large-scale networks.

We propose a novel framework leveraging the computing continuum to address these challenges by turning client heterogeneity into an advantage. By quantifying data diversity, the framework clusters similar clients, assigns them to edge servers, and trains cluster-specific models. This approach simplifies the global model training task into smaller, manageable subtasks while enabling simultaneous learning of client, cluster, and global models, thereby enhancing FL's scalability and efficiency.

## 2. Results

We propose pHFedKD, a hierarchical federated learning framework that addresses client heterogeneity, concept drift, and scalability in decentralized environments. The framework combines federated geospatial clustering (FGC), hierarchical aggregation, and multi-teacher knowledge distillation (KD) to optimize client, cluster, and global model performance.

Key contributions include:

- Dynamic Clustering: FGC uses a novel client affinity (CA) metric to adapt to spatio-temporal mobility and concept drift by dynamically reassigning clients during training.
- Hierarchical Framework: Spanning three levels (clients, edge servers, and cloud), it trains local models, aggregates cluster-specific edge models, and produces a global model.
- Multi-Teacher KD: Enables inter-cluster knowledge sharing, preserving cluster-specific traits while leveraging global insights.

Evaluated on the CityScapes dataset for semantic segmentation, pHFedKD outperformed state-of-the-art baselines like FedAvg and FedProx in both static and dynamic settings, achieving up to 19% improvement in client accuracy, 24.7% in edge accuracy, and reducing communication costs by 50% through lightweight logit transmissions.

## 3. Planned Follow-up Activities

The finished work is currently under review for publication at the IEEE International Conference on Parallel and Distributed Processing Symposium (IEEE IPDPS). In the meantime, we are focused on adapting the framework for intelligent farming, specifically for tracking animals in remote alpine regions, to validate its broader effectiveness.

## 4. Suggestions for Contributions by Third Parties

The framework developed in this project is applicable to diverse domains such as autonomous systems, smart cities, and precision agriculture, addressing critical needs for privacy, mobility, and real-time processing. It offers potential for integration with existing FL frameworks and edge infrastructures. Future studies could explore meta learning-driven clustering, privacy-preserving mechanisms like differential privacy, and resource optimization techniques to enhance adaptability, regulatory compliance, and efficiency in energy-constrained environments like IoT devices.