Alex Iacob

aai30@cam.ac.uk | **+44 7587 052460** | **GitHub** | **LinkedIn** | **Google Scholar**

Education

Oct 22–Present *PhD in Computer Science*, **University of Cambridge**
- Third-year PhD Candidate in Machine Learning Systems. **Advisor: Dr. Nicholas Lane**.
- Specialization: **Distributed Optimization** for training foundation models across geo-distributed datacenters under strict bandwidth/VRAM constraints.

Oct 21–Jul 22 *MPhil Advanced Computer Science*, **University of Cambridge**
- **Distinction (Rank 5/36, 84%)**. **Advisor: Dr. Nicholas Lane**.
- Dissertation: *The Local-Global Trade-off in Federated Learning*. Published at **EuroMLSys**.

Oct 18–Jul 21 *BSc Computer Science*, **King's College London**
- **First-Class Honours (85%)**. Recipient of the **Undergraduate Research Fellowship Award**.

Selected Publications

Oral Top 1.8% ICLR-25 **Alex Iacob**, et al. "DEPT: Decoupled Embeddings for Pre-training Language Models".

Top 3% ICLR-26 **Alex Iacob**, et al. "MT-DAO: Multi-Timescale Distributed Adaptive Optimizers with Local Updates".

Top 5 % ICLR-26 **Alex Iacob**, et al. "DES-LOC: Desynced Low Communication Adaptive Optimizers for Training Foundation Models".

Best Paper FL@FM **Alex Iacob**, et al. "Worldwide Federated Training of Language Models". Published at the Federated Learning in the Age of Foundation Models venue at *NeurIPS 2024*.

Work Experience

May. 24–Present **Research Scientist**, *Flower Labs*, Cambridge
- **Model Training:** Pre-trained **1-13B** models on **hundreds of billions of tokens** outperforming baselines such as SmolLM2 and OLMo2 in downstream tasks with matched compute.
- **Engineering:** Developed the aggregation layer for **Flower Photon**, unifying **32 H100 GPUs** across **4 geo-distributed datacenters** (US/EU) by leveraging **Local SGD**-based methods.
- **Research:** Published **DEPT**, enabling arbitrary vocabulary scaling with limited memory/bandwidth. Created two local-update adaptive optimizers, providing **convergence guarantees for heterogeneous loss functions**, established new SOTA.
- **Optimization:** Migrated pre-training codebase to **torchtitan** and **torchft**. Achieving a **4x wall-clock speedup** over baseline Python implementation.

Jan. 23–Present **Teaching Assistant**, *University of Cambridge*
- Authored the primary lab codebase for the university's **first** Federated Learning course.

Jun. 20–Oct. 22 **Undergraduate Research Fellow**, *King's College London*
- Applied computational social choice techniques to **multi-agent system decision-making**.

Technical Skills

Tools PyTorch, **torchtitan**, **torchft**, Docker, WandB, Slurm, Hydra

Distributed ML 4D Parallelism (FSDP/TP/PP/SP), Multi-Node Training, Local SGD

Research Areas Optimizer Design, Mixture-of-Experts, Distributed LLM Pre-training, Unlearning

Extended Publications

MLSys ’25	L. Sani, **Alex Iacob**, et al. “Photon: Federated LLM Pre-Training”.
NeurIPS ’25	W. F. Shen, X. Qiu, M. Kurmanji, **Alex Iacob**, et al. “LLM Unlearning via Neural Activation Redirection”.
ICLR ’26	W. Zhao, Y. Chen, W. Ma, Y. Tang, S. Hu, S. X. Hu, **Alex Iacob**, et al. “Rethinking Data Curation in LLM Training: Online Reweighting Offers Better Generalization than Offline Methods”.
NeurIPS FL@FM ’24	L. Sani, **Alex Iacob**, et al. “The Future of Large Language Model Pre-training is Federated”.
MobiUK ’23	**Alex Iacob**, et al. “Robust and private multi-modal federated human activity recognition”.
MLSys On-device ’23	**Alex Iacob**, et al. “Privacy in multimodal federated human activity recognition”. In: *3rd On-Device Intelligence Workshop*, *MLSys ’23*.