

Yanda Tao

Personal website : taoyanda.github.io

Email: yanda.tao85@gmail.com

Mobile: +33695075999

LinkedIn: yanda-tao

EDUCATION

- **Imperial College London** London, United Kingdom
MSc. in Advanced Computing Oct. 2022 - Oct. 2023
Grade: 77.66%, Distinction
- **CentraleSupélec (University of Paris-Saclay)** Gif-sur-Yvette, France
Diplôme d'ingénieur, in Computer Science and AI Sep. 2019 - Jul. 2023
Relevant courses: High-performance computing, Object-oriented Software Engineering, Cloud Computing and Distributed Systems, Optimization
Grade: 15.2/20 (Postgraduate), 13.46/20 (Undergraduate)

RESEARCH EXPERIENCE

- **LSDS, Department of Computing, Imperial College London** London, UK
Pre-doctoral Research Assistant, funded by Samsung Mar. 2024 - Present
- **Master's thesis** London, UK
Imperial College London 2023
Supervisor: Prof. Peter Pietzuch
Title: "HetML: Heterogeneity-aware Automatic Parallel Training for Distributed Deep Learning"
- **Research module** London, UK
Term-long research project Jan. 2023 - May. 2023
Advisor: Prof. Peter Pietzuch
Subject: Automatic Parallelism for Training Deep Learning Models
 - Conducted a literature study on existing distributed and parallel training systems for DL
 - Established a comprehensive topology on parallelism techniques and auto-parallel systems
 - Produced a 30-page report with relevant technical content to discuss potential research directions

PAPERS

- **HetML: Heterogeneity-Aware Auto-Parallel Training for Distributed Deep Learning**
Yanda Tao, Pedro Silvestre, Marcel Wagenlander, Peter Pietzuch, *manuscript in preparation* 2023
Draft (thesis) available.

PROFESSIONAL EXPERIENCE

- **SAP France** Levallois-Perret, France
Junior Full Stack Developer Mar. 2022 - Sep. 2022
 - **Implement integration tool:** Adapted close-sourced integration tool for GitHub project management.
 - **Implement dependency checks:** Managed patch conflicts based on commit dependencies using traversal algorithms on customized dependency graphs.
 - **Delivery:** Supported the delivery process of crucial patches for commercial tools alongside delivery managers.
- **Alstom Transport** Saint-Ouen, France
Data Scientist/MLOps intern Sep. 2021 - Feb. 2022
 - **Radio log analysis:** Analyzed 3M+ radio logs to diagnose and predict on-board radio equipment health issues.
 - **Pipeline deployment and maintenance:** Built a Cloud-based data analytic pipeline to accelerate and automate data processing, baseline creation, and prediction using OpenFaaS and Apache NiFi.
 - **Pipeline optimization:** Improved prediction speed by 10x by aggregating queries and pruning irrelevant data transfers between cloud storage and online databases.
- **ViaRézo, Computer Science and Network society of CentraleSupélec** Gif-sur-Yvette, France
Member of technical team Oct. 2019 - Feb. 2021
 - **DevOps:** Developed and maintained containerized applications on private cloud for 2000+ students.
 - **SysAdmin:** Managed the society-run private server cluster with 20+ servers with IaC tools.

SKILLS

- **Programming:** Python, Java, C++ (Intermediate)
- **Tools:** DeepSpeed, PyTorch/JAX, XLA, Kubernetes, IaC tools (Ansible/Terraform)
- **Platforms:** Deployment experiences on Azure Cloud and AWS
- **Languages:** Chinese and French native speaker, English (IELTS 8.0), Spanish (B1)