

# Ixeia Sánchez Périz

DATA & ML ENGINEER • CLOUD DATA PLATFORMS • LARGE-SCALE ML SYSTEMS

Zürich (Switzerland)

+41 78 230 9596

Mail

LinkedIn

GitHub

Scholar



## EXPERIENCE

### SWISS INTERNATIONAL AIRLINES

Jun. 2024 - present | Zürich

#### Cloud Data Engineer • Top Performer on 1st year

- Designed and **productionized terabyte-scale PySpark** pipelines, re-engineering a 15k+ LoC legacy SQL module into a modular, metadata-driven pipeline, reducing processing time **from 6h to 30m** for tabular data.
- Prototyped **agentic** agents for SQL to PySpark migration and drove AI adoption across engineering and product teams.
- Reduced **deployment** time **from 3h to 5 min** for **15 engineers**, enabling parallel releases and cutting failure hours per release by **driving** the **CI/CD design** and implementation in Microsoft Fabric and Power BI.
- Communication:** Presented our Data Platform Solution to the Fabric product team at **Microsoft** HQ in Redmond, US.

### VISIUM (CONSULTING)

Mar. 2023 - Aug. 2023 | Zürich

#### Junior Machine Learning Engineer

- Market Data:** Developed **energy hedging** method with 35% savings and constant CVaR applying forecasting & interpolation.

### CORIFY CARE (MEDTECH)

Jun. 2020 - Jul. 2021 | Valencia (Spain)

#### Data Scientist Intern (Part-Time 50%)

- Signal Processing:** real-time ML pipeline on medical devices (64 electrodes, >20 patients), achieving low-latency **inference** for ECG classification.

## EDUCATION

### ETH ZÜRICH

Sep. 2021 - Apr. 2024 | Zürich

#### MSc in Data Science

- Thesis (Honours): Sailor**, Scalable System for **LLM training in the Cloud**.
  - Developed **distributed training** pipelines for **1.5B+** parameter LLMs using PyTorch & DeepSpeed, on top of a resource planner. Built on top of Kubernetes (GKE), Docker, NCCL, Gloo.
  - Implemented a novel in-CPU checkpointing mechanism (**2.87x** higher **goodput** on preemptible VMs) with parallelism.
- Cloud Computing:** Designed dynamic Kubernetes **scheduler** achieving SLO 1.5ms 95th percentile latency.
- Computer Vision:** Domain adaptation for tree canopies segmentation from drone images to **satellite imagery**.
- Audio Data:** Automated an **ablation pipeline** to validate +250 Kaggle solutions on **self-designed** strategies.

### POLYTECHNIC UNIVERSITY OF VALENCIA (UPV)

Sep. 2016 - Jul. 2021 | Valencia (Spain)

#### Double BS in Telecommunication Engineering & Business Administration and Management

- Ranked 1st** (avg. 9.1/10, 33 courses with honours)
- Thesis: Markov model for the cost-effectiveness study for healthcare system sustainability (published in Q1 journal).
- Exchange programme at Singapore Management University: Business studies • Santander Bank Scholarship

## SKILLS

**PROGRAMMING & FOUNDATIONS** Python, SQL (advanced), Bash, Git • Data structures, parallelization, testing (pytest)

**DATA ENG & ORCHESTRATION** PySpark, Databricks, Microsoft Fabric, dbt, Fivetran  
CI/CD (Azure Pipelines, Terraform, GitHub Actions) • Workflow orchestration (**ADF**)

**ML & AI** PyTorch, DeepSpeed, Multiprocessing • LLM training, MCP, **MLOps**  
NLP (transformers, text preprocessing), Computer Vision, ECG, time-series

**CLOUD & INFRA** Azure, GCP (BigQuery, GKE), **Docker, Kubernetes** • Distributed computing (NCCL, Gloo)

**HONORS & RECOGNITION** 

---

- |  |                              |
|--|------------------------------|
| • <b>Commencement Speaker</b> for BSc graduation ceremony at University      | May 2025 • Valencia (Spain)  |
| • <b>Youngest Speaker</b> at a University Career guidance talk at Uni Zurich | Nov. 2024 • Zurich           |
| • <b>Best Student Award</b> by Social Council (500+ students)                | Jul. 2022 • Valencia (Spain) |
| • <b>Nova Top 111 University talent in Spain</b> (under 25)                  | Mar. 2022 • Madrid (Spain)   |

**LANGUAGES** 

---

**English** (C1) • **German** (B1/B2) • **French** (A1) • **Spanish** (Native)