



# Xoel Garcia Maestu

Data Science & Engineering · AI & Deep Learning

📍 A Coruña, Spain    ✉️ xoel.maestu@gmail.com    ☎️ 604 02 65 59  
 🌐 www.xoelmaestu.com    🔗 xoelmaestu    📱 xoel-maestu

## Education

- BSc** **University of A Coruña (UDC)**, Data Science and Engineering A Coruña, Spain  
Sept 2022 – July 2026
- Full data lifecycle: capture, preparation, modeling, deployment, and visualization
  - Combining Data Engineering, Statistical Analysis, and AI/Deep Learning applied to real-world use cases
- Prof. De-gree** **Conservatorio Profesional de Música de A Coruña**, Clarinet A Coruña, Spain  
Sept 2014 – June 2025
- Official regulated music education (Elementary + Professional Degrees) · wind bands, chamber music, and solo/accompanied performance

## Experience

- INDITEX**, Data & AI Engineer Arteixo, A Coruña · On-site  
May 2026 – present  
3 months
- Data Science, Data Engineering, and Artificial Intelligence within INDITEX's technology team.
- NTT DATA**, Data Analytics & BI Intern A Coruña · Hybrid  
Sept 2025 – Nov 2025  
3 months
- Technology consulting internship for the banking sector within a workforce data analytics project.
- Built **Python** ETL pipelines for data cleaning, integration, and validation across heterogeneous banking sources
  - Generated **synthetic datasets** to support historical analysis and testing where real data was unavailable
  - Delivered **Looker Studio dashboards** for business stakeholders; **automated periodic reporting workflows** via scripts
  - Migrated **Google Apps Script (JavaScript)** logic to **Python**, improving maintainability and traceability

## Skills

**Machine Learning & AI:** Scikit-learn, Imbalanced-learn, XGBoost, LightGBM, Optuna, Hyperopt, PyTorch, Keras, TensorFlow, HuggingFace (MobileNet, ResNet, YOLOv8x, DINOv2, SigLIP...), MLflow · CNNs, ResNets, DenseNets, U-Nets, Mask R-CNN, LSTMs, GRUs, VAEs, Autoencoders · Federated / Continual / Distributed learning, RL (Q-Learning, SARSA), NLP, Apache Lucene

**Data Engineering:** Apache Spark, Spark MLlib, ETL pipelines, PostgreSQL, MySQL, Oracle, MongoDB, Snowflake, PostGIS, Denodo, Docker, Git, AWS, Azure

**Data Science & Analytics:** Pandas, Polars, NumPy, PySpark, Matplotlib, Seaborn · Inference, hypothesis testing, regression, ARIMA, bootstrap · Power BI, Looker Studio, Tableau, MDX / OLAP

## Programming Languages and Tools

**Languages:** Python, R, SQL, Java, Julia, JavaScript

## Spoken Languages

**Galician:** Native

**Spanish:** Native

**English:** C1, Cambridge English Advanced (CAE)

## Projects

---

### Bachelor's Thesis: Explainable Machine Learning for Post-Liver-Transplant Survival Prediction with an Integrated Clinical Platform

July 2026

- Decision-support system that predicts **five-year** post-transplant survival from donor and recipient data, on **real clinical data from CHUAC** (258 cases): a complete, reproducible *pipeline* (MissForest imputation, feature selection, and Bayesian hyperparameter optimization) comparing six models (decision tree, random forest, XGBoost, and multilayer perceptron), with the decision tree as the transparent reference validated by a clinician (**AUC-ROC 0.757**), and a **dockerised clinical web platform (React + Django REST, PostgreSQL)** with a structured registry, assisted donor-recipient allocation, natural-language search via a local LLM (Qwen2.5 · Ollama), and global and individual explanations of the predictions (SHAP, LIME, PDP, FGClustering)
- **Grade: 10/10, Distinction (Matrícula de Honor)** · BSc in Data Science and Engineering, [UDC](#) ([IRLab](#) research group)

### MXP AutoScan: Computer Vision for Vehicle Assessment

Oct 2025 – Jan 2026

- **Dual Mask R-CNN** for automated damage detection, part identification, and severity estimation. REST API and web interface generating structured technical reports for insurers, workshops, and mobility companies
- **1st Prize, Datathon "O Camiño dos Datos" 2026** · [DIHGIGAL](#) · [ITG](#) · [UDC](#) · [IGAPE](#)

### MXNJ Outfit Intelligence: Fashion Visual Product Recognition · HackUDC 2026

Feb 2026 – Mar 2026

- Given a fashion photograph, identify every garment and return the matching product reference from a 27k-item Zara catalogue: YOLOv8x-seg anatomical segmentation + **SigLIP SO400M & DINOv2 ViT-L/14** ensemble; contrastive fine-tuning (InfoNCE + Triplet) + semi-supervised **B2B self-expansion** → **69.66% Recall@15** · 100% open-source (PyTorch, OpenCLIP, HuggingFace)
- **2nd Prize, INDITEXTECH Challenge & 1st Prize, Best Use of Open Source AI** · [HackUDC 2026](#), [GPUL](#) · [UDC](#)

### LocalFold: Full-Stack Web Platform for Protein Structure Prediction with AlphaFold2 and Generative AI · Impacthon 2026

Apr. 2026

- Dockerised web platform (**React 19 + FastAPI, Cloud Run + Firestore**) for protein structure prediction with **AlphaFold2 (CESGA FT3)** for researchers without HPC/SLURM knowledge, from a sequence or natural-language input. Mol\* 3D viewer, PAE and biological metrics; differential analysis **Gemini/Vertex AI** PDF, Telegram bot, multilingual UI · **A2A multi-agent**: orchestrator + 4 specialist agents with **Google Search Grounding** · JWT, PWA
- **2nd Place, Best Use of Google Cloud Tools**, [Impacthon 2026](#), GDG Santiago de Compostela

### BujanSat: European Space Agency (ESA) CanSat Program

June 2021 – Oct 2021

- End-to-end satellite subsystem replica (sensors, live radio telemetry, autonomous logging, parachute deployment) within a soda-can form factor; secondary mission: seed dispenser for wildfire environmental awareness
- **1st place** Galician regional + Spanish national phases → represented Spain at European final → ESA **Best Outreach Award**

More personal projects at [xoelmaestu.com](#)

## Selected Honors

---

- **1st Prize**, Datathon "O Camiño dos Datos" · [DIHGIGAL](#) · [ITG](#) · [IGAPE](#) · [UDC](#) (Jan 2026)
- **2nd Prize, INDITEXTECH Challenge** · [HackUDC 2026](#), [GPUL](#) · [UDC](#) (Feb 2026): multimodal fashion retrieval pipeline; 69.66% Recall@15 on 27k-product Zara catalogue
- **Best Use of Open Source AI Award** · [HackUDC 2026](#), [GPUL](#) · [UDC](#) (Feb 2026): 100% open-source stack (PyTorch, SigLIP, DINOv2, YOLOv8x, OpenCLIP)
- **2nd Place, Best Use of Google Cloud Tools** · [Impacthon 2026](#), GDG Santiago de Compostela (Apr 2026). LocalFold: dockerised web platform (React 19 + FastAPI, Cloud Run + Firestore) with A2A multi-agent system and Gemini/Vertex AI for AlphaFold2 protein structure prediction for researchers without HPC expertise
- **Best Outreach Award**, European CanSat Competition, [European Space Agency](#) (2021)
- **1st place**, [Galician regional](#) & [Spanish national](#) CanSat phases, ESERO Spain (2021)
- First prizes & honorable mentions: Festiciencia, Galiciencia, Maker Faire, Ciencia en Acción, OpenScience Cambre, Foro Intercomunitario de Investigación Juvenil de Vigo

## Personal Interests

---

- **Music:** avid listener and active practitioner; enthusiast of jazz, classical, electronic, experimental and all kind of music
- **Climbing & outdoors:** certified **High-Performance Athlete** (Xunta de Galicia · DOGA, Jun 2022, valid until 2026) · competitive sport climbing (bouldering, lead, speed) with Galician podiums and Spanish championship finals · **athlete representative at the Galician Mountaineering Federation (2022–2026)** · climbing livestream commentator · board member of Club de Montaña Coruña (Social Media) · hiking, outdoor climbing, surf and spending time in nature
- **Science fair volunteer and participant:** project presentation, organization support, social media management, and live interviews with researchers across national and international events
- **Travelling and meeting new people and cultures:** Erasmus+ stays in Italy and Lithuania during secondary school developing maker biotechnology projects: design and construction of low-cost lab equipment (centrifuges, spectrophotometers, incubators, magnetic stirrers...) working with international teams
- **Documentary recording** and art across audiovisual, pictorial, sculptural, installation, performance, and relational formats