

Marc Molina Van den Bosch

+34 632 645 284 | marcmolinavdb@gmail.com | github.com/molinamarcvdb | Nationality: Spain

Machine learning researcher specialized in geometric deep learning, generative modeling, and privacy-preserving methods (differential privacy, federated learning, secure aggregation). I design and ship ML systems on complex biomedical data — from graph- and transformer-based architectures for spatial reasoning over 3D structures, to diffusion models and large language models — with a strong focus on translating research into production-ready code.

EXPERIENCE

Industrial PhD Researcher in Private & Federated Machine Learning

Apr. 2025 – Present

CERN – in collaboration with Universitat Pompeu Fabra / Full-time

Geneva, Switzerland

- Designed novel second-order optimization methods for differentially private deep learning (DP-KFC, DP-Shampoo) – data-free and architecture-driven preconditioners that recover utility lost to DP-SGD across CNNs, transformers, and medical segmentation networks (ICML 2026; MICCAI 2026 under review).
- Built differentially private federated learning pipelines integrating secure aggregation and post-hoc explainability for cross-silo medical imaging; quantified the interplay between DP, client heterogeneity, and explanation fidelity (DECAF, MICCAI 2025).
- Developed geometric deep learning architectures (decoder-free supervoxel GNNs) for multi-modal 3D medical imaging applied to brain-tumor localization (GRAIL, MICCAI 2025).
- Contributing to the CAFEIN federated learning platform at CERN, translating research methods into deployable infrastructure for privacy-preserving ML across institutions.
- Working across architectures (LLMs, diffusion models, GNNs, transformers, CNNs) and modalities (imaging, sequence, tabular) in both federated and centralized settings.

Machine Learning Researcher

Feb. 2024 – Present

Kantonsspital Aarau / Full-time

Aarau, Switzerland

- Advanced differentially private federated learning across graph, convolutional, and transformer-based architectures for segmentation and classification, leveraging synthetic data generation to mitigate utility loss.
- Developed a privacy-preserving synthetic image generation and evaluation pipeline with conditional generative models (GANs, Diffusion, Latent Consistency) using PyTorch, TensorFlow, and HuggingFace.
- Implemented GitLab CI/CD for automated testing and deployment, with MongoDB for scalable metadata management and reproducibility across environments.
- Built and deployed a web platform for expert evaluations (Turing tests) used to align generative models with human judgment.

Data Science Intern

Aug. 2023 – Jan. 2024

Roche / Full-time

Basel/Kaiseraugst, Switzerland

- Developed a package for analyzing missingness, sample progression, and data quality in clinical trials.
- Application of ML models (PyTorch, TensorFlow, Keras) to analyze missing data patterns and anomaly detection in interventional, post-marketing and longitudinal clinical trials.
- Designed and implemented an end-to-end MLOps pipeline using Kedro for structured and reproducible workflows, including PySpark integration for scalable data processing. Leveraged Docker for containerization and GitLab CI/CD for continuous integration and deployment.
- Utilized explainable ML methods (SHAP) to enhance prediction modelling interpretability in clinical trials.
- Led NLP project extracting trends from social media using LLMs and Retrieval Augmented Generation systems on Azure.

Biomedical Computer Vision Developer

Jun. 2021 – Aug. 2023

Vall d'Hebron Research Institute / Part-time

Barcelona, Spain

- Built a Graph U-Net for joint node and edge classification on vessel graphs derived from volumetric segmentations — geometric deep learning over 3D anatomical structures.
- Implemented cerebral artery and large vessel occlusion segmentation with convolutional- and transformer-based architectures on CTA volumes.

- Designed an end-to-end automatic thrombus detection and segmentation pipeline from preoperative imaging.
- Developed algorithmic ML solutions supporting ischemic stroke treatment workflows.

Quality Control Intern

Pharmore Research SLU / Part-time

Sept. 2020 – Sept. 2021

Barcelona, Spain

- Performed quality control of patient and care-giver completed questionnaires.
- Managed and updated company database and server.

EDUCATION

PhD in Private Machine Learning

CERN – in collaboration with Universitat Pompeu Fabra

Geneva, Switzerland

Apr. 2025 – Present

Master of Science in Medical Informatics

Fachhochschule Nordwestschweiz

Muttenz, Switzerland

Sept. 2022 – Sept. 2024

- Thesis: "A multifaceted evaluation framework of generative medical image models"
- Best Thesis Award: "Novartis Diplompreis 2024"

Bachelor in Biomedical Engineering

EEBE - Universidad Polit cnica de Catalunya (UPC)

Barcelona, Spain

Sept. 2018 – Aug. 2022

- Thesis: "Automatic Cerebral Arteries segmentation for ischemic stroke treatment"

High School in Science

Col legi Lestonnac

Barcelona, Spain

Sept. 2016 – Jun. 2018

PUBLICATIONS & AWARDS

Paper under review at MICCAI 2026

Mar. 2026

- "Architecture-Driven Preconditioning for Sample-Efficient Differentially Private Medical Image Segmentation"

Paper at ICML 2026

Jul. 2026

- "DP-KFC: Data-Free Preconditioning for Privacy-Preserving Deep Learning"

Paper at MICCAI 2025

Sept. 2025

- "The Interplay Between Explainability and Differential Privacy in Federated Healthcare"

Paper at MICCAI 2025

Sept. 2025

- "Decoder-Free Supervoxel GNN for Accurate Brain-Tumor Localization in Multi-Modal MRI" — geometric deep learning over 3D supervoxel graphs.

Paper Presentation NeurIPS 2024 Safe Generative AI Workshop

Dec. 2024

- "Memorization Detection Benchmark for Generative Image models"

Abstract at the Radiological Society of North America (RSNA)

Sept. 2024

- "Enhancing federated learning in multicenter studies via synthetic data generation for improved downstream performance"

Best MSc thesis Award: Novartis Industry Diplompreis

Sept. 2024

- "A multifaceted evaluation framework of generative medical image models"

Paper presentation in Information Systems for Business Management

Sept. 2024

- Evolutionary, Bayesian, and Quasi-Monte Carlo Hyperparameter Tuning

Top 5 abstract award at KSA's "Tag der Innovation & Forschung".

May 2024

- Abstract on "Conditional Style-based GANs for Generating Synthetically private Mammograms"

Poster presentation as European Stroke Conference (ESOC) in Munich

May 2023

- Abstract on "Fully automated Large Vessel Occlusion localization and segmentation from CT and CTA"

LANGUAGES

English: Full professional proficiency.

Spanish and Catalan: Native.

SKILLS & INTERESTS

Skills: Python, PyTorch, TensorFlow, JAX, HuggingFace, Hydra, Weights & Biases, Geometric Deep Learning (GNNs), Generative Models (GANs, Diffusion, LLMs), Differential Privacy, Federated Learning, MLOps, Docker, GitHub, GitLab CI/CD, Apache Spark, PySpark, Kedro, MongoDB, SQL, Linux, Rust, JavaScript (React, Node, Express), MATLAB, ITK, VTK, Slicer 3D, Probability Theory, Biomedical Statistics

Interests: Football, Hiking, Reading, Swimming, Cooking and FCBarcelona