Mariano Rodriguez

Chief Scientific Officer | Applied AI Research Leader

PhD Computer Vision & Machine Learning | Speech, Vision & Language Models in Production

rdguez.mariano@gmail.com • +33 7 82 58 82 37 • Paris, France • rdguez-mariano.github.io

Applied-AI research leader who converts research into shipped, revenue-bearing product. Co-founder and CSO of **Freeda**, where I own the scientific strategy and AI roadmap and lead a team of 3 ML engineers building AI compliance review for the construction industry. Previously founding ML engineer at **Gladia** (speech-to-text API, PAYG and enterprise customers; later acquired by OVH), ML engineer at **Powder** (highlight detection in gameplay video), and sole technical founder of **Nano Quant Services**, a value-stock ML product I built and ran from zero. PhD and postdoc in computer vision from ENS Paris-Saclay; author at WACV and CVPR; 6 years turning applied research into systems that hold up under production load and customer scrutiny.

Selected Impact

- **Built a research function from scratch.** Hired and lead 3 ML engineers at Freeda; set the AI roadmap, the evaluation bar and the in-house-vs-foundation-model boundary.
- **Research that reached the P&L.** Compliance workflows, engine and infrastructure delivering the fastest verifications in the market, supporting a EUR 1M+ book of business, and underpinning a EUR 2.5M round closed with Frst and Brick & Mortar plus EUR 20M with Notion as lead investor.
- **Every layer of Freeda's AI stack shipped under my direction.** Perception models beating vendor baselines, plan description layer, agentic verification, document intelligence, codegen mini-apps, and self-hosted Triton/Celery infra.
- **Inference economics at scale.** As founding ML engineer at Gladia, rebuilt Whisper serving to lower cost per hour of audio at better WER.
- **Zero-to-one, with capital discipline.** Sole technical founder of Nano Quant Services: full product — scraping, training, assistant with hybrid search & contextual RAG, backend + frontend — shipped in 6 months.

Leadership & Experience

Co-Founder & Chief Scientific Officer
Nov 2024 – Present

Freeda, Paris

Own scientific strategy, research roadmap and the production AI platform for automated compliance review of pre-construction plans against building regulation.

- **Team & roadmap.** Built and lead a team of 3 ML engineers; own hiring and the AI roadmap.
- **Product impact.** Took the compliance workflows, engine and infrastructure from blank page to production in a few months, cutting manual plan-review time 3× with respect to the market standards.
- **Commercial results.** Live with 10+ paying customers within 10 months; the scientific work was the direct technical basis of the EUR 3.5M round raised from Frst and Brick & Mortar, and EUR 18.5M from Notion.
- **Cost & reliability.** Architected the self-hosted serving stack (GCP Compute Engine, NVIDIA Triton, Celery, LangChain, 4 DBs [Elasticsearch, Postgres, Redis, ChromaDB], monitoring [Grafana, Loki, Prometheus], GitHub CI/CD, auto-deploys); the first working prototype ran at < EUR 300/month with an NVIDIA L4 GPU.

Founder & Chief Technology Officer
May 2024 – Nov 2024

Nano Quant Services, Paris

Sole technical founder. Built and operated a complete machine-learning product for identifying value stocks, then made the call on its future.

- **Built from zero.** Shipped an end-to-end product alone in 6 months: data acquisition (Selenium), model training, RAG retrieval, FastAPI backend, React/Node frontend, Kubernetes deployment on premise.
- **Capital discipline.** Reached 2 paying users on EUR 20/month of infrastructure, zero outside capital and 4 months of runway — no outside engineering spend.

Founding Machine Learning Engineer — Speech
Jan 2023 – May 2024

Gladia, Paris

Applied research at commercial scale: speech-to-text infrastructure sold as a public API to 100+ customers (PAYG and enterprise contracts); Gladia was later acquired by OVH.

- **Inference research with a direct margin effect.** Rebuilt Whisper for production serving (ONNX, NVIDIA Triton, flame-graph profiling): lower cost per hour of audio at better WER than the market at the time.
- **Owned a research area end to end.** Speaker diarization and voice-activity detection (pyannote, silero), which unlocked a customer segment: the meeting-transcription use case.

Machine Learning Engineer — Video & Multimodal
Jul 2021 – Jan 2023

Powder, Paris

Applied computer vision inside a consumer gaming product.

- **Research shipped to consumers.** Event and highlight detection in gameplay streamer video.
- **On-device deployment.** Game identification in video and spoken-language identification exported to ONNX/CoreML.
- **Published what shipped.** Multi-annotation attention model for video summarization, CVPR Workshop 2023.

Postdoctoral Researcher, Applied Mathematics
Jul 2020 – Jul 2021

École Normale Supérieure Paris-Saclay

Industry-funded research contracts — research run to an external client's specification and deadline. Advisor: Gabriele Facciolo.

- **Huawei.** Recurrent-CNN joint video denoising and demosaicing delivered under contract; published at WACV 2023.
- **Kayrros.** CNN cloud detectors for satellite imagery.

Technical Depth

Research	Computer vision, NLP, speech recognition & diarization, agents/skills/RAG, video understanding, geometric estimation, optimisation, applied mathematics
Modelling	PyTorch, PyTorch Lightning, TensorFlow, Keras, OpenCV, scikit-learn, LangChain, ONNX, CoreML
Serving	GCP, AWS, Slurm, Kubernetes, Docker Swarm, NVIDIA Triton Inference Servers, FastAPI, Flask, flame-graph profiling
Engineering	Python, C++, ctypes (C++↔Python) and MEX (C++↔MATLAB) bindings, PyTest (fixtures & mocks), SQLAlchemy/Postgres, React & Node, GitHub CI, Cookiecutter, Task, Make & CMake, Unix administration, LaTeX/TikZ
Scientific tooling	Comet/WandB, Gin config, QGIS, MATLAB
Languages	Spanish (native), French (fluent), English (fluent)

Education

PhD, Computer Vision & Machine Learning — École Normale Supérieure Paris-Saclay. Advisors: Jean-Michel Morel & Julie Delon. 2016 – 2020

MSc "Mathématiques, Vision, Apprentissage" (MVA) — ENS Paris-Saclay. Highest Honors. 2015 – 2016

MSc 1 "Mathématiques et Modélisation" — Université Paris Cité. Highest Honors. 2014 – 2015

BSc Mathematics — Universidad de La Habana. Highest Honors; top student of the faculty. 2007 – 2011

Selected Publications

Video joint denoising and demosaicing with recurrent CNNs WACV 2023

Multi-annotation attention model for video summarization CVPR-W 2023

Robust estimation of local affine maps and its applications to image matching WACV 2020

Covering the Space of Tilts SIIMS 2018

Full list: rdguez-mariano.github.io/publications

Awards & Fellowships

Laureate, FMJH Doctoral Program — Fondation Mathématique Jacques Hadamard. Three-year competitive fellowship. 2016

Laureate, PGSM Masters Scholarship — Paris Graduate School of Mathematics. Two-year international scholarship. 2014

Top student, Faculty of Mathematics & Computer Science — Awarded by the vice-chancellor, University of Havana. 2011