Juan Gabriel Garassino

GenAI / LLM & Machine Learning Engineer

Berlin, Germany
juan.garassino@gmail.com
+49 152 24024 860
artificial-artifacts.com
linkedin.com/in/juan-garassino
github.com/juan-garassino

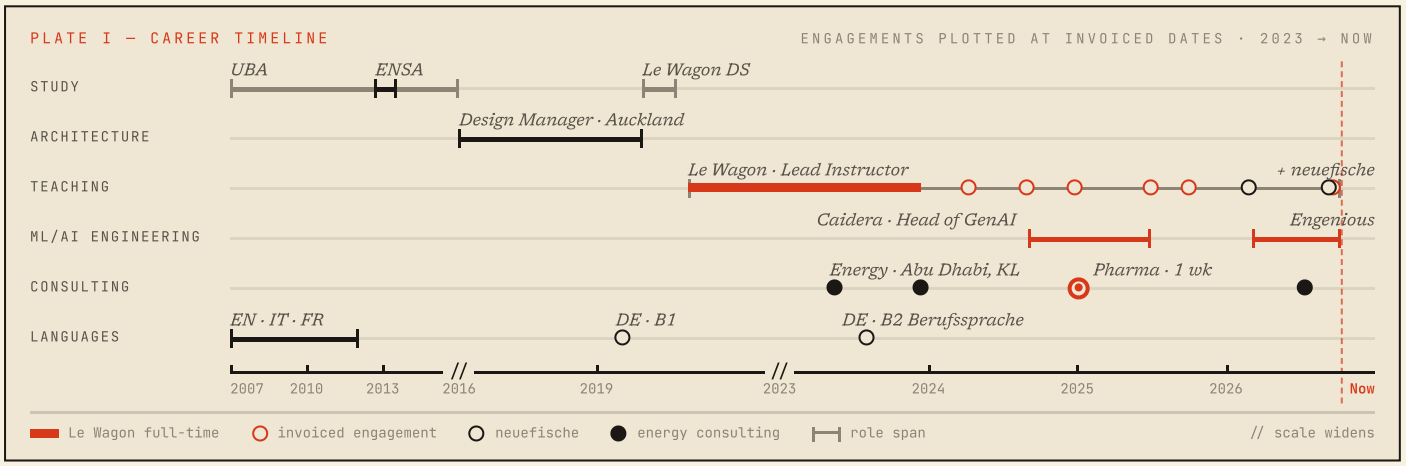
Builds production LLM and multi-agent systems, and teaches the engineers who run them.

Six years across ML and AI engineering, consulting and teaching. Builds from scratch first, on cloud and local models, and uses frameworks where they fit. Owns the full lifecycle, from design and evals to infrastructure, deployment on AWS, GCP and Azure, and monitoring. Trained as an architect.

§ 00 At a glance

BY THE NUMBERS

6+	300+	9	3 +2
years in AI & ML	*students & pros taught*	*teaching cities*	*languages, plus 2 basic*



§ 01 Experience

LEDGER, 2016 → PRESENT

01 CURRENT ROLE

Forward Deployed AI/ML Engineering Consultant

Engenious Inc · *US, remote from Berlin* — **Mar 2026 – Present**

- Production LLM systems: multi-agent pipelines as Dockerised SQS workers on AWS ECS, plus Terraform-provisioned MCP integrations.
- Lead engineer on asynchronous document AI services: OCR and page classification, LLM agents that extract and summarise key facts, and a tabular model for structured predictions.
- CI/CD with GitHub Actions, LLM-judge quality gates and adversarial testing; hands-on lead for client stakeholders and product scope.

02 TEACHING & TRAINING

Lead Data Science & AI Instructor

Le Wagon · *full-time 2021–2023, then freelance B2B / B2C* **neuefische** · *2026* — **2021 – Present**

- Taught ML, Deep Learning, LLM Engineering, MLOps and Data Engineering to 300+ students and professionals, in bootcamps and corporate cohorts.
- Delivered complete programmes end to end, from fundamentals to demo day; since 2026 also capstone mentoring and career hosting at neuefische.

CITIES Berlin · Tokyo · Amsterdam · Kuala Lumpur · London · Dubai · Munich · Barcelona · Riyadh

§ 01 Experience

CONTINUED

03 ENGINEERING LEADERSHIP

Head of Generative AI Engineering

Caidera GmbH · *Berlin, Germany*

Sep 2024 – Jun 2025

- Deployed multi-agent systems with tool calling, memory and context engineering, plus AI evaluations and monitoring.
- Architected a content and compliance AI platform end to end: a concierge agent routing to 7 production LlamaIndex workflows on Azure. Work featured by LlamaIndex.

04 ARCHITECTURE & DESIGN

Design Manager

European Stone Masons Ltd · *Auckland, New Zealand*

2016 – 2019

- Directed CAD/BIM workflows, 3D modeling and on-site coordination for civic and commercial projects, integrating parametric design and digital fabrication (Rhino + Grasshopper).

§ 02 Consulting

CLIENTS UNDER NDA

2023 · 2026 Abu Dhabi · Kuala Lumpur

AI Consultant, Energy Sector

Confidential client · *3 on-site engagements*

- Nine applied models: seismic and reservoir 3D U-Nets, a well-log transformer, diffusion inpainting, segmentation, turbofan RUL and PPE detection; Keras, MLflow, Prefect, Cloud Run.

Jan 2025 Heidelberg, Germany

AI Consultant, Pharma & Diagnostics

Confidential client · *1-week sprint*

- Feasibility analysis and architecture for internal LLM assistants and document-retrieval tools, later built by internal teams.

§ 03 Skills

AUDITED AGAINST THE CODE

MODELS	Transformers · Diffusion · GANs · VAEs · Mixture-of-Experts · Mamba · ViT · U-Net
TRAINING	RLHF · GRPO · DPO · PPO · LoRA · PEFT · Quantization · Distillation
FRAMEWORKS	Python · PyTorch · TensorFlow · scikit-learn · LlamaIndex · LangGraph · MCP · FastAPI · MLflow · deepEval
INFRASTRUCTURE	Docker · Kubernetes · Terraform · GitHub Actions · Postgres · pgvector · Redis · Neo4j · Pinecone
CLOUD	GCP · AWS · Azure · RunPod

§ 04 Projects

LIVE AT ARTIFICIAL-ARTIFACTS.COM

Artificial Artifacts

Live multi-agent portfolio assistant: SQL, vector and knowledge-graph specialists that draft résumés and job-fit analyses.

PromptPlot

Text-to-pen-plotter pipeline: an LLM writes validated G-code and the pen moves while the model is still generating.

DeepSculpt

Generative 3D voxel sculpture with a 3D GAN, diffusion and a 3D VAE, exported to OBJ/STL.

Scaffold

Conversational parametric BIM: agents plan, generate and validate building geometry on a knowledge-graph canvas.

PRINTS 37 published pen-plotter plates: ML mechanisms and the Millennium Prize problems.

§ 05 Education

2020 Data Science & Engineering · Le Wagon, Berlin

2012–13 Exchange in Parametric Design · ENSA, France

2007–16 Diploma in Architecture & Urban Design · UBA

§ 06 Languages

Spanish Native · English C1 Cambridge

German B2 Berufssprache, TELC

Italian A2 · French A2