

ANA MUÑOZ

SENIOR DATA SCIENTIST & GENERATIVE AI ENGINEER

Mexico City, Mexico | ana.clavius@gmail.com | +52 55 1152 0035 | [linkedin.com/in/anadatascientist](https://www.linkedin.com/in/anadatascientist)

PROFILE

Senior Data Scientist and Generative AI Engineer with a background in Physics and 4+ years of experience designing, building, and deploying production-grade AI/ML systems. Specialized in LLM application development—including multi-agent architectures, Retrieval-Augmented Generation (RAG), end-to-end content generation pipelines, and multimodal AI spanning text, image, and video. Strong foundation in classical machine learning: deep learning, time series forecasting, optimization, NLP, semantic search, and computer vision. Proven expertise in building scalable solutions on Google Cloud Platform (Vertex AI, BigQuery, Cloud Run) with solid MLOps practices. Experienced leader who has built an AI practice from scratch, managed multidisciplinary teams, and delivered multiple high-impact projects in parallel across retail, food, agriculture, healthcare, and education. Excellent communicator with a track record of translating complex technical concepts into clear business value for stakeholders and executives.

PROFESSIONAL EXPERIENCE

Thinkcare Technologies - Google Cloud Partner

Apr 2024 - Sep 2025

Head of Data Science & AI

Mexico City, Mexico

- Founded and led the AI and Data Science practice from the ground up: defined the solution portfolio, identified high-value use cases, designed system architectures, and built a multidisciplinary team of 5—recruiting, mentoring, and developing junior engineers to mid-level contributors.
- Architected and deployed an LLM-powered duplicate detection system integrating textual similarity (RapidFuzz), geolocation analysis, and LLM-based data standardization, with optimized large-scale parallel processing—significantly improving operational efficiency for one of Mexico's leading retailers.
- Designed and built a route optimization engine using constrained KMeans clustering and Google Maps APIs, incorporating business rules for commercial impact and strategic coverage—boosting field operations by 30%.
- Developed a LangChain-based multi-agent system for automated medical appointment scheduling, integrating Google Calendar, Gmail, and Drive APIs for end-to-end automation and confirmation tracking.
- Led internal adoption of Generative AI through targeted training programs for executives, developers, and PMs—reducing development time by 20% and timeline deviations by 15%.
- Managed strategic client and partner relationships (Google Cloud), leading technical and financial proposals, resource planning in GCP, and cross-functional project delivery.

Simón - EdTech Microlearning Platform

Oct 2025 - Dec 2025

AI Content Pipeline Engineer

Mexico City, Mexico

- Designed and implemented a full end-to-end AI content generation pipeline (syllabus → script → video) integrating multiple state-of-the-art models (Gemini 3, Nanobana, Veo 3.1) to accelerate course production by 400% compared to manual workflows.
- Applied advanced prompt engineering techniques—few-shot prompting, chain-of-thought scaffolding, and narrative consistency strategies—to ensure multi-module courses maintain coherence and professional quality throughout.
- Optimized deployment on Google Cloud Run by efficiently managing memory constraints and parallel API calls, achieving 10x faster generation speed and enabling a highly scalable production workflow.

Grupo Herdez - Multinational Food Company

Aug 2023 - Apr 2024

Data Scientist

Mexico City, Mexico

- Developed custom LLM-based systems using advanced GCP services to enable natural-language querying of BigQuery, delivering real-time insights in graphical, tabular, and descriptive formats to non-technical business users—an early production deployment of Generative AI (2023).
- Led weekly demand forecasting for major wholesale clubs (SAMS, Costco, City Club) by designing and deploying time series models (Prophet) from scratch, with recurring training and evaluation pipelines that significantly enhanced prediction accuracy.

Greenspin (AgriTech) | CAIRO Research Center

Oct 2022 - Mar 2023

Data Scientist - Master's Research Internship

Würzburg, Germany

- Developed a deep learning system for satellite-based wheat crop detection in collaboration with CAIRO and Greenspin, building convolutional neural networks from scratch in PyTorch with transfer learning to improve performance in data-scarce regions.
- Surpassed previous state-of-the-art accuracy by 10% and reduced satellite data preprocessing time by 60%. Codebase became the technical foundation for future projects at both organizations.

TECHNICAL SKILLS

Generative AI: LLM application development, Multi-agent systems (LangChain, Google ADK), Retrieval-Augmented Generation (RAG), Prompt engineering (few-shot, chain-of-thought), Multimodal AI pipelines (text, image, video), LLM-based data standardization, Semantic search, OCR with multimodal models

Machine Learning: Deep Learning (PyTorch, TensorFlow, Keras), CNNs, Transfer Learning, Time Series Forecasting (Prophet), Statistical Modeling, NLP, Computer Vision, Route Optimization, Clustering algorithms

Cloud & Engineering: Google Cloud Platform (Vertex AI, BigQuery, Cloud Run, AutoML), Docker, GitHub, CI/CD, API integration (OpenAI, Gemini, Google Maps, Calendar, Gmail, Drive), Scalable pipeline design, MLOps

Programming: Python, R, SQL, Scikit-Learn, NumPy, Pandas, Matplotlib

Analytics & Visualization: Looker Studio, Executive dashboards, KPI reporting, BigQuery analytics

CERTIFICATIONS

Google Cloud - Generative AI Leader Certification

Aug 2025

Google Cloud - Professional Machine Learning Engineer

Jul 2025

EDUCATION

Master's in Data Science

2021 - 2023

Instituto Tecnológico Autónomo de México (ITAM)

- Master's thesis internship in Germany (BAYLAT scholarship) - AI research at CAIRO & Greenspin
- Host of the podcast "Algorithms + Data + Society"

Bachelor's in Engineering Physics

2016 - 2021

Universidad Iberoamericana

- Academic exchange: INSA Lyon - Department of Bioinformatics, France | Fall 2019
- Science outreach: Weekly radio show at Ibero 90.9FM (2020-2022)
- Speaker: Biblioteca Vasconcelos, Science Communication (2019, 2021)

ADDITIONAL INFORMATION

Languages: Spanish (native), English (C1 - professional proficiency), French (B2)

Activities: Music - pianist and drummer, music production & composition