

JOSÉ ADRIÁN GARCÍA GARAVITO

AI Systems & Data Engineering Analyst

adrian.garcia.garavito@gmail.com
github.com/AdagaGG
+52 9994343881

adaga.tech · García, N.L., Mexico

Engineer building production-grade AI pipelines at the intersection of scientific computing and industrial automation. Deploys end-to-end ML systems — from OCR data extraction to Bayesian optimization engines — validated against real regulatory and materials standards. Operates at the boundary between domain science and deployable code; comfortable taking a problem from first-principles analysis to a running Streamlit/API endpoint.

TECHNICAL PROJECTS

Automated OCR Compliance Pipeline

2025 – Present

Python · pdfplumber · Pixtral Vision API · Streamlit · SQLite · REST

- Built dual-format OCR ingestion pipeline processing Mill Test Certificates against 12+ metallurgical norms (ASTM/SAE/NMX), eliminating manual QA review in materials procurement workflows.
- Implemented format-detection logic for two certificate layouts (vertical rows / horizontal rows) with automatic vision-API fallback for scanned documents.
- Deployed on Streamlit Cloud with audit trail, PDF report generation, and multi-user authentication (streamlit-authenticator).

Deployed · github.com/AdagaGG/mtc-validator

Computer Vision Segmentation System

2025 – Present

Python · PyTorch · MobileSAM · OpenCV · CUDA 12.x · ONNX Runtime

- Deployed GPU-accelerated instance segmentation pipeline (MobileSAM + OpenCV) reducing image analysis time from 45 minutes to under 5 seconds per sample on ASTM E112/G1 compliant datasets.
- Validated on hypereutectoid steel microstructures; pipeline architecture extensible to any high-contrast instance segmentation domain.
- Benchmarked inference throughput with ONNX Runtime on CUDA 12.x hardware.

Bayesian Optimization Engine

2024 – 2025

Python · XGBoost · scikit-learn · Bayesian Optimization · BoTorch

- Designed ML pipeline (XGBoost + Gaussian Process surrogate) to navigate a 3-dimensional synthesis parameter space (power, velocity, DPI) for laser-induced graphene electrodes.
- Validated optimization loop on Attempt-7 dataset (P=35–45%, V=150 mm/s, D=1000 DPI); pipeline architecture transferable to any black-box optimization problem.

Scope under revision pending additional experimental data.

NLP Compliance Automation System

Jan 2026 – Present

Python · BETO · RAG · Streamlit · n8n · SQLite

- Built hybrid NLP system combining BETO (Spanish BERT) embeddings and RAG retrieval to automate psychosocial risk assessment across 72-item survey datasets per NOM-035-STPS-2018.
- Designed n8n workflow integration for automated report delivery and compliance audit trail generation.

AI Operations Infrastructure

2026 – Present

n8n · Groq API · Webhook architecture · Gmail API · Semantic Scholar API

- Deployed production webhook-based job analysis pipeline: HTTP POST ingestion → Groq LLM (llama-3.3-70b) analysis → structured fit report (score, gaps, recommendation) delivered to Gmail.
- Built automated research monitor (Semantic Scholar API) with sequential rate-limit-aware request chaining across 5 keyword tracks.

Operates on n8n.cloud with published production URLs.

WORK EXPERIENCE

Data Automation Engineer (Freelance)

Oct 2025 – Present

- Delivered B2B data cleaning and validation pipelines for international clients; improved contact database deliverability from 50% to 97% using Python (pandas) ETL workflows.
- 5-star rating on B2B Lead Generation project; managed technical delivery under SLA constraints autonomously.

Web Operations & Digital Systems

Jun – Sep 2025

- Deployed e-commerce platform and digitized inventory database, modernizing operational data infrastructure for retail business.

EDUCATION

B.Eng. Materials Engineering (IMT)

FIME – UANL · Aug 2023 – Jun 2028 (Expected)

6th semester · GPA maintained for international mobility eligibility

AWARD

Resourcefulness Prize

Bladesmithing Competition — AIST/TMS · 2024–2025

AISI 5160 cutting tool fabrication: normalize, quench, temper. Full artisan execution.

LANGUAGES

English — C1 Advanced (EF SET, Jan 2026)

Spanish — Native

TECHNICAL SKILLS

Languages & Frameworks

Python · SQL (básico) · JavaScript (n8n)

ML / Optimization

PyTorch · XGBoost · scikit-learn · Bayesian Optimization · RAG · NLP (BETO/BERT)

Computer Vision

OpenCV · MobileSAM · ONNX Runtime · CUDA 12.x

Data Engineering

pandas · pdfplumber · SQLite · ETL pipelines · REST APIs

Infrastructure & Ops

n8n · Streamlit Cloud · Docker · Git/GitHub · VS Code + Copilot

Domain Knowledge

ASTM E112/G1 · ASTM/SAE material specs · NOM-035-STPS-2018 · Electrochemical characterization

TARGET ROLES

ML/AI Engineer (Junior)

Data Engineer / ETL Analyst

AI Operations Analyst

Materials Informatics Researcher

Mitacs / DAAD International Research

Available for remote / flexible-schedule roles only — academic mobility requirements maintained.