

Juan Daniel García Belman

AI Data Automation Engineer

danielgb331@outlook.com +52 (461) 173-3822 Celaya, Querétaro, México LinkedIn GitHub Portfolio



[View Portfolio](#)

PROFESSIONAL SUMMARY

Data Engineer with **3+ years of enterprise experience** building production ETL pipelines for Fortune 500 clients (Telcel, Citi Banamex), now specializing in **AI-powered workflow automation and intelligent agents**. Proven track record processing 5M+ daily records using Azure/Databricks architectures. Currently building **autonomous AI automation systems using n8n orchestrations, LangChain, and RAG pipelines** to transform manual processes into intelligent, self-learning workflows that drive operational efficiency and business value.

CORE TECHNICAL SKILLS

AI & Automation

n8n Workflow Orchestration • AI Agents (Agentic AI) • Python Automation • LangChain • RAG Pipelines

Data Engineering

Databricks (Lakehouse) • Medallion Architecture • PySpark • Apache Airflow • SQL (T-SQL, PostgreSQL) • Delta Lake

Cloud & Infrastructure

Azure (Synapse, Data Factory, DevOps CI/CD, Monitor) • AWS (Lambda, S3) • Docker • Git • 24/7 Monitoring

PROFESSIONAL EXPERIENCE

Freelance AI Automation Engineer

May 2025 – Present

[Independent Projects](#) | [Remote](#)

- Architected **Autonomous DataCenter AI Monitor** using n8n + AI Agents for intelligent infrastructure monitoring, implementing anomaly detection algorithms and automated alerting systems for critical uptime scenarios (99.9% availability target)
- Developed **RAG-powered Healthcare AI Assistant** with LangChain for autonomous medical knowledge retrieval, processing 10K+ document embeddings with vector similarity search for context-aware Q&A generation
- Built end-to-end automation pipelines integrating n8n workflows with Python scripts, reducing manual operational tasks by 70% through intelligent decision-making agents

ETL Process Developer (Data Operations)

May 2024 – Apr 2025

[NTT DATA – Telcel Client](#) | [Mexico City](#)

- Designed enterprise **Data Warehousing architecture** with Azure Synapse Analytics, orchestrating pipelines via Data Factory to ingest 5M+ daily telecom records from on-premise SQL Server to Azure Data Lake (Parquet format)
- Optimized **PySpark jobs on Databricks** (clusters up to 10 nodes), implementing Delta Lake ACID transactions and broadcast joins that **reduced processing time by 30%**
- Deployed **CI/CD pipelines with Azure DevOps**, using ARM Templates for Infrastructure as Code across Dev/QA/Prod environments
- Implemented **24/7 monitoring framework** with Azure Monitor + Log Analytics, resolving production incidents within SLA (<2hr response time)

Metadata Governance Analyst

Feb 2023 – May 2024

[NTT DATA – Citi Banamex Client](#) | [Mexico City](#)

- Established **data governance framework** with Azure Purview, automating metadata discovery across 800+ banking tables using Python scripts (PyODBC)
- Documented **data lineage for 150+ ETL workflows**, ensuring CNBV regulatory compliance through comprehensive technical documentation

Data/Business Analyst

Feb 2022 – Feb 2023

[NTT DATA – Citi Banamex Client](#) | [Mexico City](#)

- Designed and implemented **Python automation framework** (Pandas/pyodbc) to validate 500+ monthly data attributes, reducing validation cycles from 8 hours to 45 minutes (90% time savings)
- Created **10+ interactive Power BI dashboards** with real-time KPI tracking using DAX measures and DirectQuery connections to Azure Synapse for executive decision-making

Industrial Process Analyst

May 2018 – May 2021

[East West Industrial Engineering \(ZF TRW\)](#) | [Celaya, México](#)

- Implemented **real-time data collection system** for 12 CNC production lines, optimizing cutting parameters that increased output by 150 units/hour (5% annual cost savings)

FEATURED PROJECTS

Autonomous DataCenter AI Monitor

Intelligent infrastructure monitoring with AI agents, anomaly detection, and 24/7 uptime tracking

[View on GitHub](#)

AI Agent Healthcare Assistant

Medical Q&A platform using RAG for autonomous knowledge retrieval from clinical documents

[View on GitHub](#)

Insurance Analytics Platform

End-to-end analytics with interactive dashboards and automated data pipelines

[View on GitHub](#)

Enterprise Data Warehouse

Production warehouse processing 341K+ records with Medallion Architecture

[View on GitHub](#)

EDUCATION & CERTIFICATIONS

Education

B.Eng. in Manufacturing Technologies

Universidad Politécnica de Guanajuato, 2015

Professional Certifications

Microsoft Azure Data Engineer Associate (DP-203) — In Progress

IBM Data Warehousing Engineer Specialization — Coursera, 2024

Google Data Analytics Professional Certificate — Coursera, 2022

ETL and Data Pipelines with Airflow & Kafka — IBM/Coursera, 2023

LANGUAGES

Spanish: Native

English: Professional Working Proficiency