

# Manuel José Vargas Alvarez

Ingeniero de IA · Ingeniero de ML · Data Scientist

Buga, Colombia (Remoto) | Inglés, Español | [manuel@manuelvargas.dev](mailto:manuel@manuelvargas.dev) | +1 (805) 245-7835  
[manuelvargas.dev](https://manuelvargas.dev) | [linkedin.com/in/manuel-v-01091999](https://linkedin.com/in/manuel-v-01091999) | [github.com/ArepaMan](https://github.com/ArepaMan)

## PERFIL PROFESIONAL

Maestría en Data Science (University of Pennsylvania, GPA 3,93) con un portafolio de sistemas de IA entregados y medidos: un transformer de **52,9M de parámetros** entrenado desde cero, con repositorio público, demo en vivo y pesos publicados; un pipeline de detección de anomalías sobre **509.358** contratos públicos; y una plataforma GraphRAG en producción evaluada sobre un corpus anotado a mano. Desarrollador único de dos productos SaaS multi-tenant en producción. Formación en validación GMP (FDA 21 CFR) — control de fugas de datos, líneas base medidas, evidencia trazable. Busco roles de IA, ML y Data Science (remoto).

## HABILIDADES TÉCNICAS

**ML y Data Science:** Python, PyTorch, scikit-learn, pandas, Polars, DuckDB, arquitectura transformer, tokenización BPE, estudios de ablación, detección de fugas de datos, perplejidad, ROC-AUC, calibración, scoring de anomalías, benchmarks con gold set

**IA y LLM:** Retrieval Augmented Generation (RAG), GraphRAG, bases de datos vectoriales (Qdrant), grafos de conocimiento (FalkorDB), embeddings, recuperación híbrida, prompt engineering, agentes, guardrails (NeMo), OpenAI API, Gemini, Qwen, n8n

**Ingeniería:** FastAPI, React, Next.js, TypeScript, SQL, Supabase, PostgreSQL, row-level security, Edge Functions, Docker, Cloudflare Workers, Streamlit, Hugging Face, CUDA, Git, CI/CD

## PROYECTOS

### Sol — Transformer desde cero

2026

PyTorch, CUDA, BPE, Streamlit | [github.com/ArepaMan/sol](https://github.com/ArepaMan/sol) | [sol-52m.streamlit.app](https://sol-52m.streamlit.app)

- Construí un transformer decoder-only de **52,9M de parámetros** de extremo a extremo, sin pesos preentrenados: perplejidad de validación **3,719** (IC 95% 3,693–3,745) frente a una línea base de trigramas de 23,4 sobre el mismo set de 15.141 documentos.
- Detecté fuga entre particiones — **6.304** documentos de validación eran duplicados exactos de documentos de entrenamiento — y los descarté antes de reportar cualquier métrica.
- Medí un piso de varianza por semilla ( $4,490 \pm 0,0045$  ppl en 3 corridas idénticas) antes de interpretar ablaciones; el learning rate lo supera por  $\sim 271\times$ , la escala de datos apenas por  $\sim 18\times$ . 40k iteraciones, 1,31B tokens, 16,9 h en una sola GPU de laptop de 8 GB; **154 tests** con guardia de drift en CI.

### Elephant in the Room — Detección de riesgo en contratación pública

2026

Python, DuckDB, Polars, scikit-learn, SECOP II

- Pipeline sobre **509.358** contratos de salud colombianos (COP 59,8 billones): 26.686 señales crudas reducidas a 48 hallazgos mediante cinco compuertas de falsos positivos medidas contra tasas base por estrato.
- Diagnostiqué compuertas que suprimían detecciones reales; una regla de población mínima y un denominador corregido subieron el recall posterior a adjudicación de **60% a 100%** en 11 patrones de fraude inyectados.
- Clasificador lineal UNSPSC sobre 492.025 etiquetas declaradas — **90,8% de accuracy** frente a un 53,3% de clase mayoritaria; el modelo mide, una regla determinista con umbrales publicados decide. La verificación adversarial corre contra el export publicado, no contra el código; 176 tests, 96,7% de cobertura de joins.

## Lante — Company Brain (GraphRAG)

2025 – 2026

Python, FastAPI, Qdrant, FalkorDB, Docker

- Ingesta de WhatsApp y PDF hacia un grafo de conocimiento acotado por tenant: **F1 de entidades 0,933** (0,923–0,944 en 3 corridas de LLM) frente a una línea base heurística de 0,692 sobre un corpus de 93 filas anotado a mano, 48 de ellas negativos deliberados.
- Recuperación híbrida que fusiona vectores de Qdrant con recorrido de aristas de compromiso en FalkorDB, citando remitente y marca de tiempo y absteniéndose cuando falta evidencia. Reconstruí un gold set contaminado detrás de un test de higiene del corpus; **431** tests offline en verde.

## Bodega Boost — OS de retail

2025 – 2026

React, TypeScript, Supabase, Edge Functions | bodegaboost.com

- Plataforma multi-perfil en producción para América Latina, con seis tipos de cuenta desde una sola base de código: **163 Edge Functions, 698 migraciones SQL**, POS offline-first, comercio por WhatsApp, cobros con Creem.
- Esquema multi-tenant con row-level security, escrituras solo con service role en las tablas de entitlements y un test de drift que bloquea features de pago autoconcedidas; 54 specs de Playwright, 151 archivos de tests unitarios.

## Iter AI Parking

2026

TanStack Start, Supabase, Cloudflare Workers, Gemini | iter-ai-parking.manoloxzzz99.workers.dev

- Entregué un producto completo de cuatro portales — conductor, dueño, personal y operaciones — con reservas, cobros y modo offline en **menos de 40 horas** de construcción en solitario.
- Cero gasto en APIs de mapas mediante un RPC de bounding box más haversine en Postgres; copiloto Gemini con guardrails de NeMo; revisión de autorización sin hallazgos PO/P1 explotables en la pasada por personas.

## EXPERIENCIA PROFESIONAL

### Ingeniero de Validación — cGMP Consulting Inc.

Ene 2022 – Dic 2024

Illinois, EE. UU. | Clientes: AbbVie, Fresenius Kabi, Pfanstiehl, Uberlube

- Ejecuté IQ/OQ/PQ y Revisiones Periódicas de Validación; redacté FMEA y planes de validación basados en riesgo bajo FDA 21 CFR.
- Redacté SOPs, controles de cambio y desviaciones; apoyé el cierre de CAPA con equipos interfuncionales de QA.

## EDUCACIÓN

### University of Pennsylvania — MSE, Artificial Intelligence, Machine Learning and Data Science

En curso

Philadelphia, PA | GPA 3,93 / 4,00

### University of Wisconsin–Madison — BS, Chemical Engineering

2021

Madison, WI | GPA 3,72 / 4,00