

Luis Ángel Quiroz López

Ciencia de datos | Ingeniería de datos | IA

quirozcat@gmail.com · 55 6539 3328 · Ciudad de México

lobby9000.github.io · github.com/lobby9000 · linkedin.com/in/llqq

PERFIL PROFESIONAL

Físico y maestro en astrofísica (UNAM) que desarrolla soluciones de datos e IA con Python y SQL, desde pipelines de datos y dashboards hasta aplicaciones con LLMs y RAG. Experiencia freelance evaluando respuestas de LLMs con rúbricas de calidad para el entrenamiento de modelos (RLHF). Acostumbrado a presentar trabajo técnico ante públicos no especializados.

HABILIDADES PRINCIPALES

Programación: Python (pandas, NumPy), SQL, R, Bash, JavaScript, HTML/CSS

Ingeniería de datos y nube: PostgreSQL, modelado de datos (esquema en estrella), ETL, validación de datos, Spark, Snowflake, Azure ML, FastAPI, Flask, Git, GitHub Actions, Linux

Aprendizaje automático: scikit-learn, PyTorch, TensorFlow/Keras, inferencia estadística (máxima verosimilitud, selección de modelos, pruebas de hipótesis)

LLMs y RAG: API de Anthropic (streaming, uso de herramientas, conteo de tokens), Claude Code (modo headless, skills), ingeniería de prompts, RAG con ChromaDB y recuperación híbrida, evaluación de LLMs (rúbricas, red-teaming), RLHF, SFT y DPO

Aseguramiento de calidad: Diseño de pruebas (valores límite, partición de equivalencia, pruebas negativas), pruebas de API con Postman y Newman, pytest, validación de datos con SQL, pruebas de regresión en CI, Jira

BI y productividad: Chart.js, Excel (tablas dinámicas, validación de datos), Power BI y Looker Studio (básico), Google Workspace, Microsoft 365

IDIOMAS

Español (nativo) · Inglés (C2) · Italiano (B2) · Alemán (B1) · Portugués (B1)

EXPERIENCIA

Entrenador de IA, validación de software y NLP

2026

DataAnnotation

Remoto

- Probé LLMs no publicados en más de 10 proyectos con prompts adversariales que expusieron fallas de lógica y violaciones de políticas, y reporté cada una como resultado esperado contra resultado obtenido.
- Diseñé las rúbricas y métricas para calificar las respuestas de los modelos, y aporté retroalimentación RLHF y anotaciones de NLP para su evaluación y ajuste fino.

Especialista lingüístico y entrenador independiente de IA

2026

Meridial Marketplace

Remoto

- Revisé y corregí respuestas de LLMs en pipelines multilingües con humano en el ciclo, con retroalimentación estructurada para orientar su entrenamiento.

Especialista independiente en verificación de datos de campo

2026

Contrato independiente

Ciudad de México

- Recolecté y verifiqué datos de establecimientos comerciales mediante visitas presenciales y documentación digital, para mantener el conjunto de datos exacto y completo.

Desarrollador web independiente

2025

Ashland Chemical de México

Ciudad de México

- Construí un sitio web interactivo de presentación de productos para difusión en escuelas privadas, desde los requisitos del cliente hasta la interfaz publicada.

Asistente de investigación

2024

UNAM, con la Dra. Gabriella Piccinelli Bocchi y el Dr. Ángel Sánchez Cecilio

Ciudad de México

- Revisé la literatura de cosmología, redacté una revisión bibliográfica sobre perturbaciones cosmológicas y expuse la teoría al grupo de investigación cada semana.

Divulgador científico (voluntario)

Instituto de Física, UNAM

Ciudad de México

- Presenté investigación en cosmología ante públicos que van desde niños hasta investigadores de otras áreas, y construí una demostración con cámara en vivo que simula el paso de una onda gravitacional a través del espectador.

PROYECTOS DESTACADOS

Organismos modelo de *reward hacking* en GSM8K | *Python, PyTorch, LoRA, SFT, DPO, GRPO, pytest*
Pipeline de post-entrenamiento construido desde cero sobre Qwen2.5-0.5B que subió la exactitud en GSM8K de 34.8% a 43.7%, con corridas controladas de calificadores con huecos. Encontré que un calificador sesgado induce ecuaciones falsas en la cadena de razonamiento (en el 75.5% de las respuestas que copiaron una pista incorrecta), que evaden un monitor de consistencia aun cuando se penalizan las respuestas marcadas. github.com/lobby9000/reward-hacking-gsm8k

API de compensación por incentivos: caso de estudio de QA | *Postman, Newman, pytest, PostgreSQL, GitHub Actions*

Ciclo completo de QA sobre una API REST de comisiones de venta: 92 casos de prueba priorizados por riesgo, 8 defectos en la primera versión (2 críticos) y 750 de 750 pagos conciliados contra un modelo independiente en Excel tras las correcciones. github.com/lobby9000/incentive-comp-qa

Vigía: monitoreo de estabilidad operativa | *Java, Spring Boot, Flask, Django, PostgreSQL, MongoDB*

Servicio de ingesta, API de analítica y dashboard de objetivos de nivel de servicio sobre 1.4 millones de peticiones simuladas. Calibrar el detector de anomalías bajó las falsas alarmas de 18 a 1 en 40 corridas sin perder ningún incidente. 69 pruebas automatizadas. github.com/lobby9000/vigia

Pipeline de apuntes con RAG agéntico | *Python, Claude Code, API de Anthropic, faster-whisper, ChromaDB*

Convierte grabaciones de clase de cuatro horas en notas de Obsidian con transcripción local y Claude, y suma un agente de preguntas que cita la clase y el minuto mediante recuperación híbrida. 88 pruebas automatizadas. github.com/lobby9000/class-notes-pipeline

Analítica de logística en retail | *Python, pandas, SQL, Chart.js*

Caso de estudio sobre cumplimiento de entregas con 62,420 registros de envío depurados a 59,615 documentando cada exclusión, esquema en estrella, comparación entre pares con intervalos de Wilson y un dashboard interactivo. github.com/lobby9000/Retail-Business-Data-Cases

Herramienta de señales EMA/MACD y API | *Python, pandas, FastAPI, Excel, pytest*

Motor de señales validado contra un modelo de Excel de 3,815 fórmulas con diferencias menores a 5×10^{-13} , aplicado a 35 emisoras de la BMV y expuesto como una API REST con pruebas. github.com/lobby9000/macd-toolkit

LLM pequeño desde cero | *Python, PyTorch*

En el curso Introduction to Modern AI de CMU construí desde primeros principios un motor de autodiferenciación, un transformer de solo decodificador, un tokenizador BPE y Adam, y después entrené y ajusté el modelo con SFT, DPO y RL sobre GSM8K. lobby9000.github.io

Estructura de la Vía Láctea | *Python, astropy, scikit-learn, TOPCAT*

Cruce de Gaia DR3 con APOGEE-2 DR17 reducido de 682,284 a 281,612 estrellas mediante cortes de calidad documentados, con clasificación cinemática, grupos móviles con DBSCAN y la curva de rotación. github.com/lobby9000/Milky-Way-Structure

CNN para morfología de galaxias | *Python, TensorFlow, Keras*

5,608 imágenes de galaxias de MaNGA. Una CNN más profunda subió la exactitud en prueba de 81.0% a 86.3% y el recall de la clase minoritaria de 0.68 a 0.77, comparada contra InceptionV3. github.com/lobby9000/Galaxy-Morphology-NNClassifier

FORMACIÓN ACADÉMICA

Maestría en Astrofísica de Campos y Partículas

2026

Instituto de Astronomía e Instituto de Física, UNAM

Ciudad de México

- Tesis: *Ecuación de estado efectiva y producción gravitacional de partículas en modelos inflacionarios*, dirigida por el Dr. Marcos A. García García.

Licenciatura en Física

Ago 2019 – Dic 2023

Universidad Nacional Autónoma de México (UNAM)

Ciudad de México

- Titulación por alto desempeño académico, con mención honorífica.

CURSOS Y CERTIFICACIONES

- **Introduction to Modern AI**, Carnegie Mellon University (en línea), ene – abr 2026
- **Diplomado en Finanzas**, UNAM, mar 2026 – actualidad
- **zero2hero** (2026): Claude Essentials · Claude Architect · Gestión de proyectos y agentes de IA con Claude · Desarrollo de software con Python e IA (24 sesiones) · FastAPI y SQLAlchemy · Algoritmos financieros con Python · Estadística, series de tiempo y econometría con R