

JAVIER HENGDA WU LIN

Barcelona, Spain · +34 617 864 435 · hello@javierhengda.dev
javierhengda.dev · [GitLab](#) · [LinkedIn](#)

SUMMARY

AI Software Engineer. Sole owner of a production ML service at a fintech startup: rebuilt a transaction-categorization system end to end on AWS — architecture, pipeline, evaluation, infrastructure, CI/CD. B.Sc. Computer Engineering (Dec 2026). Focus: LLM systems in production, MLOps, and the measurement discipline that keeps them honest. Case studies for every project below at javierhengda.dev.

EXPERIENCE

SPRK Software & Tech

Barcelona, Spain (hybrid)
Mar 2026 – Present

ML Engineer — sole owner of the AI categorization service

- Rebuilt an inherited prototype into a production service: **4-package** Python monorepo (~**60k LOC**, **1:1** test ratio, **1,517 commits**) on AWS — ECS Fargate, CDK infrastructure as code, RDS PostgreSQL, Redis, Bedrock — with **39 RFCs** documenting every architecture decision.
- Designed an **11-step** rule-first categorization cascade (rules as versioned data: **458** merchants, MCC ranges, knowledge base) ending in a single grounded LLM call with the right to abstain — replacing a 3-provider consensus cascade after a measured correlated-errors analysis.
- Built reproducible evaluation: **7,584-row** judged corpus, deterministic replay reproducing **1,624** live model verdicts byte-identically at zero cost; measurement exposed a small-corpus illusion (0.93 → 0.523) and quantified the improvement ceiling per cause.
- GDPR-grade engineering: PII encryption with key rotation, HMAC-verified per-user isolation with row-level security in depth, erasure as processing block; hermetic CI (network-banned unit tests) with a real-infrastructure merge gate.

SELECTED PROJECTS

tun-vpn — a VPN from first principles (*Go · Linux TUN · TLS 1.3 · Docker*)

2026

- Built a layer-3 VPN in **560 lines** of stdlib-only Go: TUN device via raw `TUNSETIFF` ioctl (C-ABI-exact structs), length-framed IP over CA-pinned TLS, goroutine relay with first-error-wins semantics.
- Reproducible Docker topology proves the properties: dual `tcpdump` inner/outer capture, break-and-recover with ~**2 s** automatic reconnect, bridge isolation making the tunnel the only path.

metrics-server-dashboard — self-hosted Mission Control (*Go · WebSockets · Alpine · ApexCharts*)

2025

- Spec-first build (architecture doc + wire protocol before code): **1 Hz** hub fan-out over WebSockets with drop-on-slow-consumer delivery and memory bounded at **every** layer (client queues → archiver → store → charts).
- Embeddable agent SDK with **zero** non-stdlib imports; adaptive chart density (line → per-core heatmap past 5 cores).

SignalML — arXiv ingestion & LLM classification pipeline (*Supabase · pg_cron · OpenAI*)

2026

- Operations-first design: run-log state machines written before external I/O, per-paper failure isolation, dry-run modes, Vault-resolved cron secrets, strict structured outputs double-validated, per-run cached-token cost accounting.

Kofre — personal & social banking backend (*Java 21 · Spring Boot · GCP · Terraform*)

2025

- Team university project: **50+ endpoint** REST API (Clean Architecture, DDD); CI/CD with Cloud Build to Cloud Run (**sub-3-min** zero-downtime deploys); **5 Terraform modules** cutting provisioning to **<10 min**.

EDUCATION

University of Barcelona

B.Sc. Computer Engineering

Barcelona, Spain
Sep 2022 – Dec 2026

- Relevant coursework: Machine Learning, Computer Vision, Reinforcement Learning, Distributed Systems, Computer Security, Graphics & Data Visualization, Databases.

TECHNICAL SKILLS

Languages: Python, Go, TypeScript/JavaScript, Java, SQL, C/C++
AI/ML: LLM systems in production (Bedrock, OpenAI, structured outputs, grounding & abstention), evaluation harnesses & deterministic replay, XGBoost, embeddings
Backend: FastAPI, Spring Boot, Node.js, REST, gRPC, WebSockets
Cloud & DevOps: AWS (ECS Fargate, CDK, RDS, IAM, Bedrock), GCP (Cloud Run, Cloud Build), Docker, Terraform, GitHub Actions, GitLab CI
Data: PostgreSQL, Redis, Supabase, Firestore, SQLite
Practices: TDD, RFC/ADR decision records, hermetic testing, Clean Architecture, CI/CD, observability (Honeycomb, Sentry)
Spoken: Spanish (Native), English (Fluent), Chinese (Conversational), Catalan (Fluent)

CERTIFICATIONS

Mathematics for Machine Learning (<i>Imperial College London</i>)	May 2024
IBM Data Analyst Professional Certificate	Apr 2024