

Angelo Lacson *Senior Software Engineer — ML & AI Systems*

✉ angelo.lacson.work0407@gmail.com

🌐 linkedin.com/in/angelo-lacson-81808b3b6

☎ +63 969 051 3850, +1 209 273 2545,

📍 Santo Tomas, Pampanga, Philippines

🔗 github.com/angelo-lacson

PROFILE

ML Platform and AI Systems Engineer with 6 years of experience building production infrastructure for machine learning — from real-time feature pipelines and model serving systems to LLM-powered document intelligence platforms. Core expertise in Python, FastAPI, Spark, Snowflake, Kafka, PostgreSQL, pgvector, Redis, AWS, Docker, and Kubernetes. Delivered ML infrastructure across fraud detection, credit risk, compliance automation, and recommendation systems. Comfortable owning backend architecture end-to-end, including cloud deployment, observability, and cross-functional delivery with product and data science teams.

PROFESSIONAL EXPERIENCE

Software Engineer (ML Platform), Tecton 🌐

08/2024 – 01/2026 | CA, US (Remote)

- Worked on Tecton's ML feature platform, building backend systems around Feature Views, Aggregation Engine, and Embedding Engine to turn raw customer data into reliable features for real-time prediction and offline training.
- Improved the Aggregation Engine for high-volume batch and streaming features using Python and Spark, helping scale feature computation to 50M+ daily events and updates while reducing repeated recomputation across common time-window aggregations.
- Optimized low-latency feature serving through Feature Services and the online retrieval path, improving batching, caching, and timeout handling to increase serving throughput by about 2.5x while keeping p95 feature response time under 100ms.
- Built embedding feature workflows through Tecton's Embedding Engine, supporting text-based features for search, ranking, recommendation, and personalization use cases, and improving embedding generation reliability through better batching and failure recovery.
- Standardized 20+ production feature definitions across batch, streaming, and on-demand workflows, improving feature reuse and reducing new model integration time from several days to less than one day.
- Added production observability with OpenTelemetry and Grafana, including freshness checks, p95/p99 latency dashboards, failure alerts, and trace IDs, cutting debugging time by roughly 40% during feature computation and serving issues.

Software Engineer (AI Applied), Gatekeeper 🌐

09/2022 – 07/2024 | London, UK (Remote)

- Built Vendor Risk Analysis AI, an AI-powered third-party risk platform that automated vendor security assessments — reducing a review process that previously took most of a working day down to roughly 15 minutes per vendor.
- Designed and shipped Python and FastAPI backend services for document processing, vendor analysis, compliance mapping, remediation workflows, audit logs, and natural-language Q&A.
- Built the AI document intelligence pipeline for SOC 2 reports, ISO 27001 evidence, SIG questionnaires, and penetration test reports — using semantic chunking, embeddings, pgvector search, LLM analysis, and citation mapping.
- Implemented multi-framework compliance analysis across 12 security frameworks and 2,500+ controls, allowing a single uploaded document to be mapped against multiple standards simultaneously instead of requiring manual per-framework review.
- Delivered operational dashboards in React and TypeScript giving teams real-time visibility into vendor risk scores, compliance coverage, severity trends, and remediation SLAs.
- Hardened production readiness with PostgreSQL and pgvector data models, Redis-backed async processing, S3 document storage, RBAC, rate limiting, Docker deployment, and automated test coverage across 100+ API endpoints.

Full-Stack Developer, Self-Employed 🌐

03/2020 – 07/2022 | Remote

- Developed patient-facing web features for StoryMD's digital health platform using React and TypeScript, building health journey pages, medical content displays, and personalised user workflows — delivering a responsive and accessible experience across a platform handling sensitive medical data for real users.
- Built reusable React components and integrated Node.js and Python backend services via REST APIs, connecting frontend health record views, biomarker displays, and content navigation to AWS serverless infrastructure (Lambda, S3, SQS) handling live health data and medical content pipelines.
- Built an Education CRM platform with FastAPI and React covering student records, course management, staff roles, scheduling, and role-based access control — delivering a full administrative backend and responsive frontend for managing day-to-day institutional operations.

EDUCATION

University of the Philippines Diliman | B.S. in Computer Science 

08/2014 – 07/2018 | Quezon City, PH

Thesis: Development of a Multithreaded File Processing System Using C++ and POSIX Threads

SKILLS

- Languages: Python, TypeScript, JavaScript, SQL, Bash
- Backend: FastAPI, Django, Node.js, NestJS, REST APIs, GraphQL, async processing, microservices
- ML Platform: Tecton, Spark, Snowflake, feature stores, batch/streaming pipelines, model serving support
- AI/LLM: RAG, embeddings, pgvector, vector search, LLM APIs, document intelligence, citation mapping
- Data: PostgreSQL, Redis, DynamoDB, MySQL, MongoDB, S3
- Cloud/DevOps: AWS, Docker, Kubernetes, GitHub Actions, Terraform basics
- Observability: OpenTelemetry, Prometheus, Grafana, structured logging, distributed tracing
- Security: RBAC, audit logging, rate limiting, SOC 2, ISO 27001, vendor risk workflows

CERTIFICATES

- AWS Certified Solutions Architect – Professional — Jun 2021
- AWS Certified Machine Learning – Specialty — Aug 2023