CHRISTIAN MACION AI Engineer · Multi-Agent LLM Systems · Eval-Harness & Cost-Aware LLM Deployment

christianmacion26@gmail.com +63-991-616-2630 Digos City, Davao del Sur, Philippines (UTC+8) LinkedIn

FIT FOR: AI Engineer · LLM Application Engineer · Eval Engineer · Agent Reliability · Applied AI : Anthropic · OpenAI · Scale AI · Surge

SUMMARY

I build production multi-agent LLM systems that turn multi-day expert work into reliable, cost-aware automation : and keep them trustworthy with evaluation harnesses, model-routing cost discipline, and honest validation. I designed one reusable architecture from the ground up (no framework) and shipped it across four systems (quant research, venture incubation, content, knowledge base); I authored the curriculum that teaches others to optimize their own.

In the most intensive engagement, I operated an 11-agent orchestrator-worker research platform with contracted hand-offs, separation-of-duties between generation and validation agents, a 31-gate statistical evaluation harness implemented scipy-free in numpy, a tiered model-routing layer (Opus / Sonnet / Haiku / Fable) with per-dispatch token budgeting, structured-output contracts gated by a mechanical validator, persistent agent memory with provenance governance, and ~76,500 LOC of light-dependency Python supporting it : plus a telemetry dashboard reading LLM session logs and an offline notebook server.

Specialist in the scarce 2026 skills: eval design, agent reliability, and LLM cost modeling. Strong command of multi-agent orchestration (orchestrator-worker / supervisor-worker), tool use / function calling, MCP, Claude Agent SDK / Claude Code, RAG with measured retrieval quality, context engineering, prompt caching, observability, and structured-output validation. Comfortable across Python (numpy, pandas, pyarrow), SQL, AWS S3, Docker, and PDF/PPTX/DOCX generation. I build with large language models; I do not train them.

Before the AI work, a self-directed quant-trading practice (crypto and futures) that gave me the adversarial-validation discipline underpinning every eval gate I write. Available 30 hrs/wk for an AI / agent engineer role.

CORE SKILLS

- Multi-Agent Orchestration (orchestrator-worker / supervisor-worker)
- LLM Evaluation Harness Design (statistical-gate validation)
- Model Routing & Token-Cost Optimization
- Context Engineering
- Structured-Output Contracts & Validation (must exit 0)
- Tool Use / Function Calling
- Hand-off Contracts & Context Isolation
- Model Context Protocol (MCP)
- Claude Agent SDK / Claude Code
- LLM-as-Judge Pipelines
- Adversarial Verification & Refutation
- Regression / Holdout Discipline
- Persistent Agent Memory & Provenance Governance
- Prompt Caching
- RAG / Chunking / Source-Grounded Generation
- Retrieval Scorecards (recall@K, MRR@K)
- Observability & Telemetry Dashboards
- Python (numpy, pandas, pyarrow, boto3)
- SQL
- AWS S3 Data Lakes
- Docker (working)
- JavaScript
- PDF / PPTX / DOCX Generation
- Headless-Chrome Rendering
- AI-Assisted Research & Paper Summarization
- Statistical & Mathematical Modeling (foundational)
- AI Tools (ChatGPT, Gemini, Claude)
- Research Documentation & Communication

EDUCATION

Financial Management Studies (current) — University of Mindanao (UM) (1st Semester, AY 2026-27 (enrolled))

Engineering Studies (units completed) — University of Southeastern Philippines (USeP) (2022 - 2024)

Philippine Science High School (PSHS) — Philippine Science High School System : Southern Mindanao Campus (Graduated 2022)

Certified Technical Analyst Program — Society of Technical Analysts of the Philippines (STA Tier-1) (Dec 2025)

TECHNICAL TOOLS

- Python (numpy / pandas / pyarrow / boto3)
- JavaScript
- SQL
- AWS S3
- Docker (working)
- Google Colab / Jupyter
- Git & GitHub
- Notion
- ChatGPT / Claude / Gemini APIs
- Claude Agent SDK / MCP
- Headless-Chrome Rendering
- Microsoft Excel / Google Sheets

RELEVANT EXPERIENCE

AI Systems Engineer · LLM Research Platform (contract) 03/2026 - 06/2026

systematic-strategy desk (NDA-protected; closed past contract 03/2026 - 06/2026) — Remote

- Architected and operated an 11-agent orchestrator-worker AI research platform (~27,500 words of role-scoped agent charters) with contracted hand-off packets, few-shot routing, separation-of-duties between generation and validation agents, and a scribe-owned single source of truth.
- Designed a 31-gate statistical evaluation harness (block-bootstrap CIs, random-timing nulls, walk-forward, 5-era stability, Deflated Sharpe, CSCV-based PBO, Minimum Backtest Length) implemented scipy-free in numpy : kills false results before they cost anything.

- Built a tiered model-routing policy (Opus = judgment / Sonnet = assembly / Haiku = mechanical / Fable = hardest autonomous) with per-dispatch token budgeting : the cost layer that keeps a multi-agent system (measured at ~15× single-agent token cost) economical to run continuously.
- Engineered structured-output contracts gated by a mechanical validator (must exit 0), persistent file-based agent memory (59 typed, cross-linked Markdown records with provenance governance), and ~76,500 LOC of light-dependency Python including a telemetry dashboard reading LLM session logs.
- Caught and documented 3 of my own false positives in a single session : banking each as enforced methodology : and authored an institutional operating standard and 16 research studies adopted across the desk.

Founder & AI Systems Engineer 2026 - Present

Macion Ventures — Remote

- Built a 7-agent venture-incubation pipeline (5 judgment-tier + 2 mechanical) with 10 lifecycle skills; produced 31 decision-grade artifacts and engineered an anti-self-approval governance pattern (the agent that proposes never approves).
- Encoded Philippine tax/regulatory rules (DTI/SEC/BIR/LGU, ₱3M VAT threshold, 8%-flat vs graduated election) directly into agent and skill prompts.

AI Systems Engineer (Independent) 2026 - Present

Editorial / content automation — Remote

- Built an 8-agent content-production pipeline (topic-scout → researcher → drafter → editor → producer → art-director → chart-maker → exporter) and a 290-line AI-output (slop) evaluator scoring drafts on 13 literature-grounded metrics; drove a real draft from HEAVY (index 81) to CLEAN (index 3).
- Engineered a dependency-free rendering pipeline (HTML/SVG → headless-Chrome PNG; Markdown → publish-ready PDF).

Trading Platform Testing & AI-Workflow Research Analyst (contract) 04/2026 - 06/2026

AI-engineering consultancy (closed past contract) — Remote

- Conducted structured testing of trading platform onboarding flows to evaluate execution clarity, workflow logic, and system usability from a trader-first perspective.
- Applied AI prompting tools (ChatGPT, Claude, Gemini) to accelerate literature review and pattern recognition while manually verifying accuracy of all outputs.
- Translated execution observations into structured insights to support product and trading workflow optimization.

Independent AI Systems Engineer 01/2025 - 05/2026

Personal Portfolio & Self-Directed Study — Remote

- Designed and shipped multi-agent LLM systems end-to-end: agent charters, eval harnesses, model-routing policy, structured-output contracts, persistent memory, and Python tooling : for content automation, venture incubation, and personal-knowledge workflows.
- Built, audited, and open-sourced a portfolio of runnable AI projects (RAG scorecard, ReAct tool-calling agent with OTel traces, LLM-as-judge validated vs humans, MCP eval server, self-critiquing reflection agent, AI-slop evaluation gate).

Financial Market Educator & AI Integration Specialist (ongoing side) 12/2024 - Ongoing

Independent / Various Universities and Communities — Remote / Davao Region

- Delivered workshops, webinars, and presentations on financial markets, AI applications, and emerging technologies.
- Served as a guest speaker at universities and professional communities on topics involving AI, financial literacy, and blockchain technologies.

ACADEMIC & PERSONAL PROJECTS

- **Multi-Agent AI Architecture (Field Guide) —** Curriculum authored from first principles (no framework) covering orchestrator-worker topologies, eval-harness design, model-routing policy, structured-output contracts, persistent agent memory, and cost discipline. Adopted as the teaching reference for a research-desk engagement. *(available on request)*
- **RAG Retrieval Scorecard (rag-recall) —** Measured RAG with a retrieval scorecard: recall@3 = 0.886, MRR@3 = 0.805. Supports recall@K / MRR@K evaluation across chunking strategies and vector-store choices. *(available on request)*
- **LLM-as-Judge Validated vs Humans —** Judge harness validated vs human raters: Cohen's κ = 0.58, 95% bootstrap CI, position-bias measured. *(available on request)*

CERTIFICATIONS

AI, Data Science & Technology:

Introduction to AI Agents : DataCamp, AI Fundamentals : DataCamp, Artificial Intelligence Fundamentals : IBM, Introduction to AI : Google, AI for the Modern Workforce : Ateneo de Davao University

Finance, Trading & Economics:

Certified Technical Analyst Program : Society of Technical Analysts of the Philippines (Tier-1), Financial Trading in Python : DataCamp, Math for Finance Professionals : DataCamp, Excel for Financial Planning and Analysis (FP&A) : LinkedIn, Foundations of Growth Equity : Goldman Sachs, JPMorgan Chase Investment Banking Job Simulation : Forage

Mathematics, Statistics & Programming:

Getting Started with Python for Finance : LinkedIn, Python Essentials 1 : Cisco Networking Academy, SQL for Finance Professionals : LinkedIn, Understanding Data Science : DataCamp

Selected Events, Hackathons & Programs:

Galactic Problem Solver : NASA Space Apps Challenge (Zurich, CH) : Oct 2025, BIDA META AICCELERATE 2025 (5-day intensive) : Meta & Bayan Academy : Dec 2025, UP Data Science Society : Alternative Classroom Learning Experience 2026 : Feb 2026, LimitlessBiz: AIM ASEAN Training Programme : Limitless Lab : Mar 2026, Guest Speaker : USeP CBA Annual Business Expo 2026 : Feb 2026, Sui Builders Program Davao : Sui Foundation : Apr 2026