SUMMARY

AI engineer specializing in multi-agent orchestration, LLMOps, and Python-based quantitative systems. ~1.5 years deploying and operating LLM-powered services in production — including a multi-agent OCR platform digitizing Thai handwritten supply-chain documents (PydanticAI orchestrator, Grok 4 / Gemini / GLM workers, running 24/7 on Alibaba Cloud SAS with autonomous retry, fallback routing, and schema-validated outputs). Two years of hands-on international logistics operations provide deep domain knowledge that informs the AI systems I build. Bilingual (Thai / professional English), comfortable working async with distributed startup teams.

CORE TECHNICAL STACK

Languages	Python (primary), JavaScript / TypeScript (working), SQL
AI / ML	PydanticAI, multi-agent orchestration, LLM API integration (Grok 4, Gemini, GLM, Claude, OpenAI), TensorFlow, PyTorch, RAG patterns, structured output / JSON schema enforcement
LLMOps / MLOps	Production LLM deployment (~1.5y), multi-model routing by cost / complexity, schema-validated agent I/O, autonomous retry & fallback logic, 24/7 headless services, hybrid-cloud resilience patterns, AI-safety guardrails (Financial Harm policy layer)
Backend	FastAPI, async Python, REST APIs, Pandas / NumPy, schema validation with Pydantic
Cloud / DevOps	Alibaba Cloud SAS (production 24/7), AWS, Docker, headless service deployment
Quant / Fintech	MetaTrader 5 (Python bridge), backtesting engines, TensorFlow signal models, risk controls, CQF / ARPM frameworks
Domain	Document intelligence (OCR), supply-chain optimization, international logistics workflows, algorithmic trading

SELECTED PROJECTS

Rasy OCR — Multi-Agent Document Intelligence Pipeline

AI Systems Architect & Lead Engineer · Solo build, deployed to production

Stack: Python, PydanticAI, Grok 4, Gemini, GLM, Alibaba Cloud SAS, JSON Schema validation

- Designed a **master/worker agent architecture** with PydanticAI as the central orchestrator, routing tasks across specialized worker agents (Grok 4 for handwriting, Gemini and GLM for typed text and layout reconstruction) based on document complexity and cost.
- Performs **semantic document understanding** — analyzes spatial relationships between text, tables, and dense schematics; rebuilds relational data across multi-page documents into strictly-typed JSON Schema; resolves character ambiguities (e.g. 'O' vs '0' in serial numbers) by cross-referencing visual context.
- Enforced **strict schema validation** end-to-end so every extracted field lands as typed JSON ready for downstream ERP / customs systems — no human cleanup required.
- Deployed as a **24/7 headless service on Alibaba Cloud Simple Application Server**, with autonomous retry, model-level fallback routing, health monitoring, and continuous uptime — operated in production for ~1.5 years.
- Executed a **"Shadow Run" beta test against 1,000 unique handwriting samples**, refining the Human-in-the-loop fallback process and task-routing accuracy before production.
- **Outcome**: reduced document processing time on equivalent workloads from days to hours; unlocked digitization of handwritten Thai documents that prior OCR vendors could not read.

Algorithmic Trading Engine (Python ↔ MetaTrader 5)

Quant Developer · Personal / research project

Stack: Python, TensorFlow, MetaTrader 5, NumPy / Pandas, custom backtesting engine

- Engineered the **Python-to-MT5 execution layer** so strategies validated in backtests flow directly into live order placement, with risk controls (position sizing, max drawdown, kill-switch) baked in at the bridge level.
- Implemented and backtested strategies based on **AlphaTrend, Supertrend, and custom ML signals** using TensorFlow models trained on historical tick data.
- Worked with leveraged ETF instruments (e.g. TQQQ) on broker platforms; strategies grounded in **CQF (Certificate in Quantitative Finance) and ARPM (Advanced Risk and Portfolio Management)** frameworks.
- **Outcome**: a reusable backtest-to-live pipeline where any new strategy plugs into the same execution layer without rewriting infrastructure.

- Models the fulfillment network as a **graph** with attributes as nodes and transit lanes as edges; operational using LP / MIP solvers and Graph Neural Networks for non-linear constraints.
- Surfaces **Pareto-optimal trade-offs** across cost, lead time, and sustainability — giving operators real decision frontiers instead of single-metric optimizations.
- **Deliverable**: reference architecture combining real-world supply-chain operations knowledge with mathematical optimization — bridging a gap that pure-ML engineers cannot model without the physical-process experience.

Hybrid AI Infrastructure — Sovereign / Resilient Deployment Strategy

Infrastructure Architect · Applied in Rasy OCR production deployment

Stack: AWS / Azure (Western), Alibaba Cloud SAS (Eastern), openPangu and other open-weight self-hosted LLMs

- Architected **hybrid AI infrastructures diversifying across Western ecosystems (AWS / Azure) and Eastern ecosystems (Alibaba SAS)**, integrating open-weight initiatives like openPangu alongside self-hosted models for API-restriction resilience and sovereign-data compliance.
- Implemented **AI-safety guardrails** for financial-domain agents (Financial Harm policy layer) so quantitative AI advisory outputs stay within compliant, mathematically verifiable boundaries.
- Approach proven in production via the Rasy OCR deployment running 24/7 on Alibaba SAS with Western-model fallback paths.

PROFESSIONAL EXPERIENCE

Geodis — Operations & Document Department (Export / Import)

Logistics Operations Specialist · 2 years · Bangkok, Thailand

- Re-engineered SOPs and workload distribution across the team, **reducing customs declaration processing time from 1–3 days to ~24 hours.**
- Mediated and resolved persistent cross-border documentation disputes between the Thai BPO division and international branches, restoring operational fluidity.
- Managed high-volume bilingual (Thai / English) shipment documentation and cross-border data validation.
- Domain knowledge from this role directly informs the AI / supply-chain systems I now build.

Precious House — Project Operations Lead

Operations & Workflow Architect · Zero-to-one build

- Built end-to-end project operations from a zero state — defined SOPs, capacity models, and vendor / procurement workflows where none existed.
- Owned the full operational lifecycle from inception to fulfillment; iterated continuously using structured root-cause analysis to remove bottlenecks.
- This builder mindset translates directly to greenfield startup engineering work: comfortable shipping when specs are ambiguous and infrastructure is incomplete.

EDUCATION & LANGUAGES

Education	B.B.A. in International Business Management — Assumption University (English curriculum)
Languages	Thai (native) · English (professional, bilingual work environments)
Continuing Ed	Active self-directed study in LLMOps & Agentic AI Operations: production patterns for multi-agent systems (PydanticAI, tool-calling, schema-enforced I/O), LLM deployment & observability (routing, retry / fallback, cost-aware model selection, evaluation harnesses), and DevOps foundations (Docker, CI/CD patterns, infrastructure-as-code concepts) — with a secondary track in quantitative finance (CQF / ARPM curricula) supporting the trading / risk projects listed above.

WORKING STYLE

Async-friendly across timezones; comfortable with Slack / Linear / GitHub workflows. Spec-driven, code-review oriented, proactive on blockers. Strong written and verbal English — able to discuss architecture trade-offs directly with non-technical founders and technical leads alike. Available for full-time engagements.