KHOMSAN PHONSAI

AI Systems Engineer · Full-Stack Developer · Optimization Specialist

+66 81-558-6566 · kphonsai@icloud.com · github.com/greefeet · Khon Kaen, Thailand · cv.nectarserve.com

PROFILE

AI Systems Engineer with 7+ years of production experience building high-performance, data-intensive systems. Currently architecting **nectarserve** — an agentic AI platform with a formal execution model designed from first principles, with auditability and safety as core architecture concerns, not afterthoughts.

Background in Structural Engineering and Computational Optimization (M.Eng, KKU) provides a rigorous, first-principles approach to systems design. Experienced with local LLM inference, vector embedding pipelines, and Claude Code as primary development tool. Brings both the theoretical depth to design correct systems and the engineering practice to ship them.

TECHNICAL SKILLS

Languages	Rust · Python · C# / .NET · SQL · TypeScript
AI / ML	Local LLM inference · Vector embeddings · RAG pipelines · Anthropic API · Claude Code (primary, 6+ months) · LLM output validation · Hallucination mitigation · Structured output parsing
AI Safety	Auditable agent execution · Provenance tracking · Containment design · Formal execution modeling
Data	PostgreSQL · TimescaleDB · Redis · pgvector
Infrastructure	Docker · REST API design · Git
Design & Build	Webflow · Figma · Photoshop · Full-stack web (C# + .NET)
Foundations	Structural Optimization · Computational Engineering · Algorithm design · Systems architecture

FEATURED PROJECT

nectarserve 2024 – present

Agentic AI platform · Rust · Continuum execution model · pgvector · Multi-provider LLM

Status: Core engine functional — entering integration testing. Business layer (pairing, metering, billing) in progress.

Model Coverage

Model-agnostic by design — integrates with 6 providers via OpenAI-compatible API: **OpenAI, Anthropic, Google (Gemini), DeepSeek, Typhoon, and local LLM servers**. Provider switching requires no changes to core pipeline.

Continuum — Formal Execution Model

Designed a layered execution model **(Axioms → Theorems → Policy → Unknown)** governing how agents reason, commit actions, and manage memory. Security constraints encoded at the axiom level — not added as features after the fact.

- **Blast Radius** — three-tier containment model; bounds impact of any agent action before commit — a security constraint in the execution model itself, not a runtime check
- **Provenance** — append-only audit trail of every agent action, decision, and state transition; designed for full inspectability and forensic review
- **WorldState** — separates live facts from distilled memory, preventing conflation common in most agent frameworks
- **10-gate pipeline** — strict ordered execution: noise → thread resolve → intent → clarity → idle → mandate → constraints → blast radius → ack strategy → resolver

Practical LLM Engineering

- **LLM Output Control** — enforced structured parsing (REPLY/KNOW delimiters) so raw LLM text never reaches downstream systems unchecked; LLM output is always validated before acting
- **Hallucination Containment** — mandate-gated and blast-radius-gated execution; hallucinated actions targeting external systems are blocked before commit, not logged after
- **Provider API Adaptation** — handles per-provider differences in auth, request format, and response parsing at the integration layer; core pipeline remains provider-agnostic

Memory Architecture

Multi-tier memory system **(L0 Signals → L1 Observations → L2 Knowledge → L3 Patterns)** with Provenance and WorldState as separate stores. L2 uses exponential decay weighting (weight = $e^{-\lambda t}$) for knowledge relevance over time.

Inference & Embedding

- Integrates with local LLM servers (Ollama, candle-based) via OpenAI-compatible API — supports on-premise / air-gapped enterprise deployment
- Vector search via pgvector with HNSW index, BGE-M3 1024-dim embeddings for agent memory
- Test suite covering core pipeline gates, memory stores, and intent classification

Built entirely using **Claude Code** as primary development tool — managed via `CLAUDE.md`, memory files, and test-before-fix discipline across long agentic sessions.

WORK EXPERIENCE

Full-Stack Developer 2017 – 2024

C# / .NET · PostgreSQL · TimescaleDB · Redis · Docker

- Built and maintained production backend systems in continuous operation over 7 years — handling high-volume time-series data with TimescaleDB and Redis, with high availability throughout
- Owned full-stack delivery across C# / .NET, PostgreSQL, and Docker — owning the full lifecycle from schema design through containerized deployment
- Engineered optimization algorithms from first principles (informed by M.Eng background) to reduce data processing overhead in high-frequency pipelines
- Maintained API integrations and source-control practices across multi-year, multi-phase engagements — preserving consistency through requirement changes

Web Developer 2016 – 2017

C# / .NET · Figma · Photoshop · Webflow

- Delivered client-facing web applications with a focus on UI design and frontend–backend integration
- Worked across design tools (Figma, Photoshop) and Webflow for visual and interactive deliverables

OTHER PROJECTS

Truss Optimization — M.Eng Thesis 2016

Structural optimization · Computational Engineering · Khon Kaen University

"Automatic Design of Truss without Ground Structure" — optimization algorithm for structural form-finding without predefined topology. Mathematical constraint satisfaction methods directly inform current work on agent execution models.

Neural Network for Slope Stability — B.Eng Thesis 2012

Neural networks · Civil Engineering · Khon Kaen University

"Slope Stability Analysis with Neural Network" — early applied ML work predicting geotechnical stability from engineering parameters.

EDUCATION

Master of Engineering — Structural Engineering 2012 – 2016
Khon Kaen University, Thailand

Bachelor of Engineering — Civil Engineering 2007 – 2012
Khon Kaen University, Thailand

PROFILES

GitHub github.com/greefeet · Kaggle kaggle.com/khomsanphonsai · LinkedIn linkedin.com/in/khomsan-phonsai-41a0a0354