Thadoe Hein

AI Engineer

thadoe.work@gmail.com | thadoe.dev | linkedin.com/in/thadoe | github.com/MyBoringFacts

Profile

AI Engineer bridging research and production, specializing in MLOps, LLM systems, and scalable AI platforms. Experienced in end-to-end ML pipelines, production-grade RAG, multi-agent workflows, model lifecycle management, CI/CD, and AI service deployment with monitoring, logging, and performance optimization.

Education

Assumption University of Thailand — Bangkok, Thailand

Bachelor of Science in Computer Sciences — *GPA: 3.7/4.0*

Experience

AI Engineer — Jun 2026 - Present

CB BANK PCL

- Design enterprise RAG and AI assistant solutions; translate business needs into technical requirements and coordinate vendor implementation.
- Develop Python and SQL identity-resolution workflows using data normalization, rule-based matching, and name similarity, with traceable evidence and review flags.
- Evaluate AI responses and support user acceptance testing to identify quality gaps and validate business requirements.

AI Developer (Freelance) — Feb 2026 - May 2026

SoeMindAI, Inc.

- Improved 3 AI products across EdTech and recruitment, owning delivery from architecture to deployment.
- Cut AI costs by ~30% through prompt design and model selection; automated presentation creation from a multi-hour process to minutes.
- Improved a study platform with 10 tools, including flashcards, summaries, mind maps, and practice generation. Improved large-document answers through retrieval and reduced unnecessary token usage.
- Engineered a job preparation platform combining ATS compatibility, job matching, interview preparation, and multi-source discovery to reduce user friction and resume tailoring time.

Curriculum Designer & AI Integrator — Mar 2026 - May 2026

Kiddee Lab — *Thailand*

- Designed Intro to AI and Machine Learning curricula for youth and trained teaching staff through hands-on sessions for scalable delivery.
- Consulted on and implemented a receipt digitization pipeline using Gemini vision models and prompt engineering to convert paper receipts into digital records.
- Delivered AI literacy and everyday-use workshops for students, and AI business-strategy workshops for adult professionals.

Machine Learning Engineer Intern — Jun 2025 - Aug 2025

Language Understanding Laboratory (LU Lab)

- Fine-tuned mT5 transformer models and ran structured experiments to improve EM/F1 scores.
- Built reproducible Hugging Face pipelines for training, evaluation, and preprocessing; conducted quantitative research for model optimization and validation.

Experience Continued

AI Engineer Apprenticeship Jan 2024 - Oct 2025

Intelligent Systems Laboratory (ISL) - Assumption University of Thailand

- Designed and deployed production-grade RAG systems to reduce hallucinations and improve response reliability.
- Built end-to-end ML pipelines for data ingestion, embeddings, retrieval, and serving; operationalized LLM applications through REST APIs for real-world workflows.
- Monitored model performance, latency, and token usage for reliability; collaborated across teams to integrate AI into existing applications.

Projects

Video Intelligence & Face Search Platform Oct 2025

- Built video person-search using face similarity and embedding-based retrieval; reduced processing time by 90% through optimized detection, tracking, and embedding pipelines.
- Achieved sub-second identity search across large datasets with Qdrant. Built distributed video processing on AWS Lambda, ECS, and RDS, with a Next.js interface for timestamped results.

RAG Platform for Document Intelligence Feb 2025

- Built a document chatbot using RAG to improve answer accuracy and grounding through document retrieval.
- Reduced latency with automated ingestion and indexing; improved answer quality through ranking strategies and prompt optimization.

End-to-End OCR & Vision Pipeline for Receipt Data Localization Feb 2025

- Collected and prepared a custom receipt dataset for a computer vision pipeline to detect and localize text.
- Saved 50+ hours of manual annotation using Azure AI Vision to generate JSON labels and convert them to YOLO annotations. Fine-tuned YOLOv11 and achieved 0.93 text detection accuracy on Thai receipt images.

LLM Fine Tuning for News Summarization Aug 2025

- Fine-tuned FLAN T5 on a news dataset using a reproducible pipeline, improving summarization performance measured by evaluation metrics.
- Used structured hyperparameter tuning and dataset preparation experiments to achieve incremental quality gains.

Technical Skills

AI Systems & Machine Learning: Large Language Models (LLMs), Retrieval Augmented Generation (RAG), AI Agents, NLP, Computer Vision, Embeddings, Predictive Modeling, OCR, Model Evaluation and Prompt Optimization
Programming: Python, Node.js, Express, React, Next.js, TypeScript, SQL, REST APIs
LLM Frameworks & AI Tooling: LangChain, LangGraph, OpenAI API, Hugging Face, Vector Databases (Qdrant), Multi-Agent Systems
Infrastructure & DevOps: Docker, CI/CD, AWS, GCP

Languages

English: Professional | Burmese: Native