

Jason Phan

📍 Vietnam ✉ jason.phan.hcm@gmail.com ☎ 0382 631 367 🌐 sunfoxy2k 💼 jason-phan-450190217

Summary

Full Stack Engineer with 6+ years owning end-to-end product delivery across regulated enterprise and banking environments, including AI products, operational dashboards, workflow automation, data extraction, compliance-oriented review cycles, and production support for internal business users.

Excellent written and verbal English communication in global engineering teams, with proven success leading cross-regional delivery across Vietnam, the US, and Singapore while mentoring developers, aligning stakeholders, and translating business requirements into reliable, maintainable software and AI products in production.

Skills

AI Orchestration: LangGraph, LiteLLM, LangSmith, Pinecone

Cloud LLM and Document AI: Azure OpenAI, AWS Bedrock, OpenAI API, Azure Document Intelligence, Amazon Textract

Backend: FastAPI, Pydantic, Python, Express.js, Node.js, Docker

Frontend: Next.js, React, TypeScript, TanStack Query, Zustand, Tailwind CSS, Jest

Cloud and Data: Azure Blob Storage, Azure VNet, Power Automate, Amazon S3, AWS Lambda, Amazon API Gateway, Amazon SQS, PostgreSQL, MySQL, DynamoDB, WebSocket

Experience

Scale AI, AI Engineer

Remote

Nov 2024 – present

- Built AI agent evaluation workflows with skill-based tool execution and stateful routing for Vietnamese coding-locale frontier model benchmarks.
- Owned verifier and judge workflow stages, pass/fail gates, and regression cycles for instruction-following and tool-use evaluation programs.
- Defined rubric-driven review flows so agent outputs were validated before high-volume production evaluation runs.
- Supported deployment, debugging, and iteration of eval agent programs with global reviewer and engineering stakeholders in English.

Finoptics LLC, Lead IT Consultant

Hybrid

June 2021 – Nov 2024

- Delivered five internal AI-powered applications for banking clients, including document processing, analyst dashboards, and workflow automation for regulated internal users.
- Led Vietnam-based engineering teams while coordinating requirements, code reviews, and delivery with US and Singapore stakeholders in English.
- Drove end-to-end delivery across document AI, transaction monitoring support, and operational tooling in compliance-oriented banking environments.
- Mentored developers and aligned business stakeholders on scope, timelines, and production support for internal banking programs.

Entana Inc., Software Engineer

Remote

Dec 2020 – June 2021

- Designed supplier network dashboards with interactive graph views, filterable metadata tables, and responsive layouts for supply-chain analysts.
- Optimized Express.js and MySQL backend queries and API pagination to reduce supplier search latency and speed up dashboard load times.
- Refactored React data-fetch patterns and component structure to improve cross-browser performance for daily operational users.

Upwork, Full-Stack Developer

Remote

Mar 2019 – Dec 2020

- Top 10% Upwork Freelancer | 100% Job Success Rate | Top-Rated
- Delivered analytics, IoT, and automation products for US and international clients from requirements through production launch.
- Built data extraction and reporting solutions for e-commerce and media analytics use cases.
- Shipped real-time dashboard experiences for smart-device and talent analytics startups.

Projects

Apron (Vietnamese Locale LLM Eval Harness)

Dec 2025 – present

- Tech stack: LangGraph, FastAPI, Next.js, TanStack Query, Zustand, LiteLLM, LangSmith, AWS Bedrock.
- Built LangGraph evaluation workflows for Vietnamese coding-locale benchmarks with stateful routing, pass/fail gates, and regression batches for instruction-following and tool-use scenarios.
- Routed LLM judge and verifier calls through LiteLLM with Azure OpenAI and AWS Bedrock fallback for rubric scoring and structured output validation.
- Instrumented LangGraph runs with LangSmith tracing for node-level debug, evidence linkage, and replay of failed benchmark cases.

Hopper (Vietnamese Locale Benchmark Fixtures) • Tech stack: Python, FastAPI, schema validation, fixture libraries. • Built Python ingestion and normalization pipelines that turn industry code samples into structured benchmark fixture libraries for Vietnamese coding-locale eval runs. • Implemented schema validation and reference checks that block bad fixtures before regression suites, reducing false failures in tool-use and instruction-following benchmarks. • Versioned corpus updates with idempotent ingest handlers so benchmark runners could pin stable fixture sets per regression baseline.	Jan 2025 – Nov 2025
AI Document Translation Dashboard - Korean Development Bank • Tech stack: LiteLLM, Azure OpenAI, Pinecone, Azure Document Intelligence, FastAPI, Power Automate. • Benchmarked Korean document OCR across Amazon Textract and Azure Document Intelligence on pilot corpora in Amazon S3 before productionizing extract on Azure VNet. • Shipped a regulated banking document platform on Azure VNet: ingest mixed PDFs/Word, OCR/extract, Korean back-office QA (Phase 1), LLM translation to English (Phase 2), and secure export with Postgres audit trails. • Used Pinecone RAG to retrieve relevant approved banking document chunks per translation block; LiteLLM routed LLM calls with glossary enforcement and translation memory for terminology consistency. • Orchestrated document intake, review handoffs, and export notifications with Power Automate flows integrated into the bilingual QA workflow across 10,000+ document pages.	Dec 2023 – Nov 2024
AML Transaction Monitoring Tuning - Shinhan Bank America • Tech stack: Node.js, React, PostgreSQL, Jest, Tailwind CSS. • Replicated AML scenario logic in Node.js services, performed transaction data profiling, and validated monitoring accuracy with Jest unit tests against core banking integrations. • Built React analyst dashboard views on PostgreSQL-backed APIs for alert review, scenario tuning, and operational visibility. • Automated backtesting workflows and alert-quality validation checks to support regulated AML monitoring releases.	Jan 2022 – Dec 2023
Financial Report Analysis Automation - Citibank • Tech stack: Node.js, Python, PostgreSQL, OCR pipelines. • Built OCR automation with Node.js, Python, and PostgreSQL to ingest and store parsed financial report data for internal systems. • Ran scheduled ingestion jobs and validation rules through PostgreSQL-backed Node.js services and Python OCR pipelines. • Exposed parsed report outputs through REST and webhook integrations for downstream analytics and reporting tools.	June 2021 – Jan 2022
Real-Time IoT and Analytics Applications • Tech stack: React, Node.js, WebSocket, DynamoDB. • Built real-time IoT dashboards for smart-device startups using WebSocket event streams and live React state updates. • Delivered YouTube talent analytics and e-commerce crawling platforms with Node.js services, MySQL storage, and React reporting views.	Mar 2019 – Dec 2020
Education	
Ho Chi Minh University of Technology , Bachelor in Computer Science	Sept 2018 – May 2022