

**Job Title:** AI Engineer

**Company Summary:**

ELGO AI is a fast-growing startup redefining how enterprises build and deploy AI. We offer a powerful no-code GenAI platform that enables organizations to create custom LLM-powered applications such as chatbots, smart document workflows, and automation agents 10x faster than traditional methods.

Our mission is to simplify enterprise AI adoption by making advanced GenAI capabilities accessible to non-technical teams. We are passionate about workflow orchestration, context-aware agents, and enterprise-grade deployment, all grounded in responsible, secure, and user-centric design. Our vision is to become the go-to platform for practical AI solutions in Asia Pacific.

**Job Description:**

As an AI Engineer, you'll play a key role in the development and optimization of ELGO's no-code platform and LLM-based workflows. Instead of model training or fine-tuning, your focus will be on building reusable, configurable, and scalable LLM-powered components that can be orchestrated by users without writing code.

You'll work closely with our CTO and engineering team to implement advanced LLM workflows, multi-agent logic, evaluation mechanisms, and performance monitoring that power real-world business applications.

**Roles & Responsibilities:**

- Develop and optimize backend systems for LLM workflow orchestration (retrieval, chaining, evaluation, memory).
- Implement reusable components (e.g., retrieval nodes, validation agents, evaluators) to power the no-code platform.
- Integrate and manage third-party models (e.g., OpenAI, Anthropic, AWS Bedrock, open-source models).
- Develop and maintain multi-agent and tool-augmented AI workflows for common enterprise use cases.
- Work on RAG pipelines, contextual memory management, and prompt engineering best practices.
- Collaborate with front-end and platform engineers to support the no-code interface.
- Conduct systematic evaluations and performance tracking of LLM applications in production.
- Stay updated with emerging GenAI and agentic framework trends.

**Requirements:**

Core Requirements

- Bachelor's degree in Computer Science, AI/ML Engineering, or equivalent experience with about 2-4 years relevant working experience.
- Strong Python programming skills with experience in LangChain, FastAPI, or equivalent frameworks.
- Experience in building production-grade RAG applications and integrating with APIs.
- Familiarity with databases (e.g., PostgreSQL, MongoDB, Redis).
- Experience with cloud platforms (primary: AWS; secondary: Azure, GCP).
- Understanding of LLM architecture and practical applications.

### Key Traits

- Strong problem-solving and systems thinking.
- Bias for action and comfort working in a fast-paced startup environment.
- Clear communication and collaborative mindset.
- Curiosity and openness to new ideas and technologies.
- Willingness to contribute beyond code – including product feedback, customer empathy, and roadmap planning.

### Optional Requirements

- Familiarity with vector databases (e.g., Qdrant, Weaviate etc.).
- Experience building agentic workflows or using frameworks like AutoGen, LangFlow, LangChain or CrewAI.
- Frontend familiarity (e.g., React.js) or backend scripting with Golang.
- Prior work with evaluation frameworks for LLM performance and safety.

### **What We Offer:**

- A front-row seat to building the next generation of enterprise AI tooling.
- The opportunity to help shape one of Asia's most promising GenAI startups.
- Exposure to real-world challenges and impactful deployments.
- A supportive and transparent team environment that values outcomes over hours.
- Growth opportunities across product, engineering leadership, and R&D.

ELGO AI is committed to equal employment opportunities and welcomes applications from all qualified candidates.