# Founding AI Engineer (Agents & Prompting)

Bangalore · Full-time · In-office · 2–8 years experience

Own dynamic prompting, agent workflows, and the reliability systems that make Aurora trustworthy at scale.

## ABOUT THE ROLE

AuroraX is building Google for the offline world. Most commerce still happens in local stores and services with no live pricing, availability, or inventory online. Our consumer voice AI calls local sellers on a buyer's behalf, gets real answers, and returns structured, verified results. An invite-only beta is live, with the US as our first market.

We are a stealth-stage, pre-seed funded team, founded by ex-YC founders, with the founding team split across the US and India. You would own every AI system outside the live voice call: the agents, prompts, and evals that turn raw conversations into verified results — roughly 80% prompt and agent engineering, with the remaining 20% backend engineering.

## WHAT YOU WILL OWN

- The entire pipeline of LLMs and the harnesses around them that powers Aurora AI across the stack — from conversing with users to orchestrating the search or talking to sellers.
- Crafting, iterating on, and maintaining prompts and model configurations for LLM usage across the platform.
- Setting up observability systems and the corresponding evals to measure and improve reliability.
- Model choices and tradeoffs: picking and swapping providers based on quality, latency, and cost, with fine-tuning where it actually pays off — including self-deployment if necessary.
- Reliability and production scale for agent systems — guardrails, observability, and graceful degradation.

## YOU ARE A FIT IF

- You have first-hand experience assembling and deploying a harness for AI agents in production.
- You understand the nuances of what LLMs can and cannot do — you understand what context engineering means.
- You have a data-driven approach to model and prompt evaluation instead of vibes.
- You possess backend engineering knowledge, since building AI agents is not limited to prompting.
- You have experience with Python, FastAPI, Langfuse (or any LLM observability platform), and Postgres.

## NICE TO HAVE

- You have fine-tuned or custom-trained your own models and can accurately predict when that beats prompting a frontier model.
- You understand LLM runtimes like vLLM.
- You understand vector databases.
- You have deployed models in Azure Foundry, Amazon Bedrock, or similar.

- You understand DAGs and durable workflows.
- IIT/BITS/NIT graduate preferred.

## PROBABLY NOT A FIT IF

- You think all problems that occur in AI systems are linked to wrong prompts.
- You come from ML research with no record of shipping to live users.
- You have only built thin wrappers over model APIs, with no evals or ownership of quality.
- You can only design but not execute the changes yourself — we need individual contributors, so you need the requisite backend expertise to execute.

## COMPENSATION

Competitive cash plus a meaningful founding-team ESOP grant. Final band depends on experience and evaluation.

## HOW TO APPLY

Send a short note on a prompting or agent system you have owned: the scale it ran at, how prompts were assembled and tested, and how you measured quality. Applications go to join@aurorax.co.