

We're looking for a truly exceptional AI software engineer — the kind who's cracked the code on building real, autonomous, agentic AI applications. Think beyond basic prompt engineering. We're talking about systems that can reason, plan, take action, write and debug code, use tools, integrate APIs, and learn continuously — systems like Cursor, Devin, or AutoGPT—but production-grade. You'll own the end-to-end development of these intelligent systems. You'll work at the intersection of LLMs, reasoning engines, memory, RAG, vector databases, tool orchestration, and scalable backend architecture. This role is for someone obsessed with pushing the limits of what AI can do, not just what it can say.

Responsibilities

- Architect, build, and scale agentic AI applications using foundation models (OpenAI, Claude, etc.)
- Design systems that combine RAG, tool use, long-term memory, planning, and multi-step execution
- Develop and integrate with APIs, dev tools, and data systems (e.g., GitHub, Jira, Slack, CRMs, databases)
- Build custom orchestration layers for multi-agent collaboration, continuous learning, and task decomposition
- Optimize performance for latency, cost, and reliability in production
- Evaluate, fine-tune, and prompt-engineer foundation models for specific tasks
- Stay at the bleeding edge of LLM capabilities, agent frameworks, embeddings, and architecture patterns
- Collaborate closely with product, design, and domain experts to build magical AI-first user experiences

Who You Are

- You've already built a real, working agentic system — not just a GPT wrapper
- Fluent in Python, TypeScript, LangChain, LLM APIs, and vector databases like Pinecone, Weaviate, etc.
- Deep understanding of transformer models, tokenization, and how to push LLMs to their limits
- Experience building backend systems and scalable cloud infrastructure (e.g., AWS, GCP, Docker)
- Ability to architect complex asynchronous workflows and memory systems

You think like a hacker — fast, efficient, obsessed with real-world performance. Ideally contributed to open-source agentic projects or published notable AI experiments.

Bonus: experience with real-time collaborative tools, browser automation, or dev tool integrations.

Why Join Us

- Work with founders and engineers who deeply understand AI
- Build products that will change how people work, create, and think
- We ship fast, think long-term, and reward deep technical excellence
- Join at the ground floor of a fast-moving, AI-native startup with massive upside