

Multi-step research agent

Jordi Hako Nintcheu · 2026-07-15

A LangGraph agent that plans its own searches, checks whether it has enough to answer, and loops back if not.

Problem

Recruiters and hiring managers evaluating AI engineers increasingly screen specifically for agent-building skill, not just prompt engineering or RAG. Most demos labeled "agent" are really a single API call dressed up as one. This project shows the real differentiator: reasoning that adapts based on what it actually finds, instead of following a fixed pipeline regardless of what comes back.

Approach

A LangGraph state machine with four nodes: plan, search, evaluate, and synthesize. The evaluate step routes conditionally, back to plan if the findings aren't good enough yet (capped at three tries), or forward to synthesize once they are.

Anthropic Claude handles the planning, evaluating, and answering. Tavily runs the actual searches, built specifically for agent tool-calling rather than a general-purpose search API bolted on by hand. It's a CLI tool, not a deployed service, a single question can take 20 to 60 seconds across multiple search and model calls, well past what a serverless function's timeout allows, and there's no visitor-facing case that needs it live anyway.

Result

Proven on a genuinely tricky test question, current, with contradictory sources in the wild. The agent correctly judged its first search insufficient, searched again with a better query, and synthesized an accurate, cited answer on the second pass. Stateless between separate runs. If it hits the three-search cap without ever being fully satisfied, it still answers with what it has rather than flagging that it wasn't confident, that caveat isn't surfaced yet.

What I'd do differently

Surface it explicitly when the iteration cap gets hit without the agent ever being satisfied, right now that case looks identical to a confident answer. Also would have smoke-tested the structured output shape against the real model earlier, a quick test would have caught a reliability issue (mixing one long free-text field with a structured list in the same call) before it showed up mid-demo question.

Part of an ongoing series of applied AI and engineering projects. Full write-up and live links at jordihako.com.