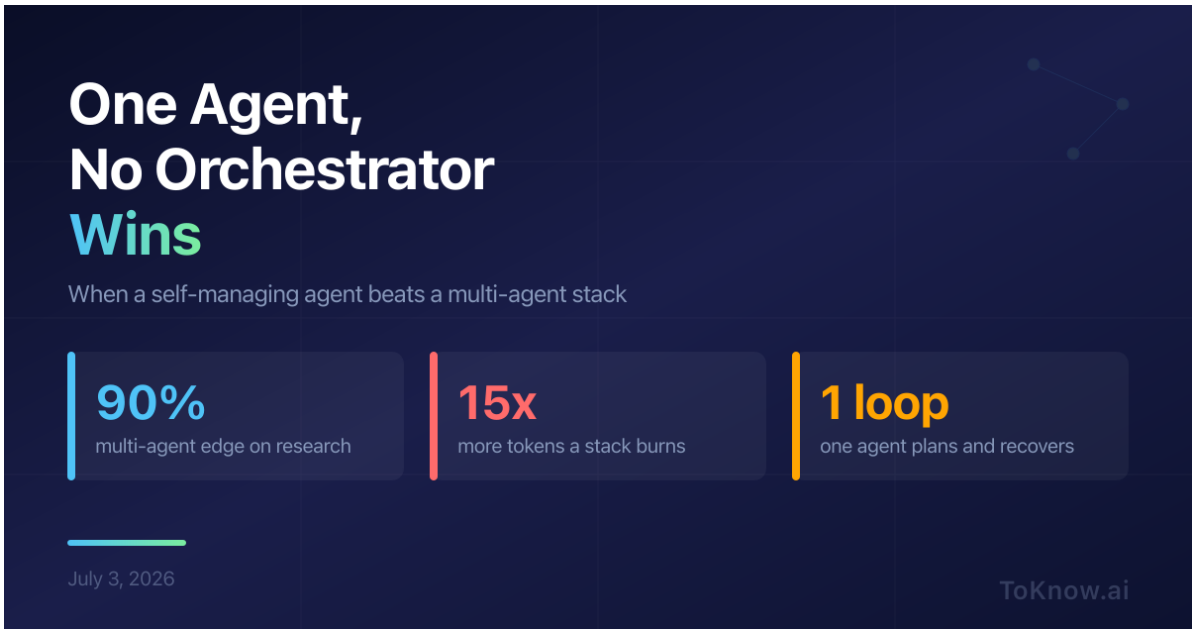
When One Self-Managing AI Agent Beats a Multi-Agent Stack

Kabui, Charles

2026-07-03

Read at ToKnow.ai



An AI agent with no orchestration layer is a single model that plans, calls its own tools, checks results, and recovers from errors, with no separate coordinator splitting the job across subagents. Anthropic draws a clean line: a workflow follows fixed, pre-written code paths, while an agent directs its own process. Cognition, the team behind the Devin coding assistant, argues in "Don't Build Multi-Agents" that for work that must stay coherent over a long run,

this single-threaded design is the safe default. Split the task across parallel subagents and each acts on assumptions the others cannot see, so their outputs conflict and the final step inherits the mess.

One agent does not always win. Anthropic found the reverse for open-ended research: a lead agent running parallel subagents beat a single agent by about 90% on its internal eval, because the work splits into independent searches that never need to agree. The catch is cost. A single agent already uses about 4 times the tokens of a chat, and the multi-agent version about 15 times as many. Anthropic also warns that most coding is a poor fit, since subtasks depend on each other and share context. A separate study across nine benchmarks found multi-agent debate often lost to a single agent while spending more compute.

So the real question is not scale but whether the work truly breaks into independent parts. Add an orchestration layer when you need real parallelism, tasks that overflow one context window, or many specialized tools. Keep the single agent when steps depend on each other, when cost matters, or when you just need something simple and reliable.

Read More: why agent teams rarely beat a single agent, and a framework that lets agents share latent state instead of text.

Sources:

- Cognition: Don't Build Multi-Agents (Walden Yan)
- Anthropic: How we built our multi-agent research system
- Anthropic: Building effective agents
- Zhang et al., Stop Overvaluing Multi-Agent Debate (arXiv:2502.08788)
