

# Production Fan-Out Patterns and a 280kLOC WebKit PR

Coding Agents Alpha Tracker

2026-06-21

## Production Fan-Out Patterns and a 280kLOC WebKit PR

*By Coding Agents Alpha Tracker • June 21, 2026*

Today's brief focuses on a concrete subagent workflow from Cognition: decompose work, keep contexts small, front-load clarifications, and let agents write prompts and sanity tests. It also flags a 280kLOC AI-generated WebKit PR as a useful case study in reviewing large agent-generated changes.

### TOP SIGNAL

The most practical signal today is **agent fan-out** from inside Cognition: a lead Devin breaks a problem into independent chunks, spins up 5-100 child Devins in parallel, then combines the results [1]. The rationale is simple and portable: agents do better when both the task and the context are small, and separate VMs make the parallelism real rather than cosmetic [1]. This is already used by Cognition's model research and product teams, not just as a demo pattern [1].

### TRY THIS

- **Use a coordinator/worker split for migrations and large refactors.** Ask the parent agent to decompose the job into independent workstreams, spawn one child per workstream, and merge centrally at the end. Concrete example: one Cognition workflow split a React Native-to-Swift migration into 6 pieces and ran them in parallel [1].
- **Make the parent agent write the child prompts.** Instead of hand-writing every worker brief, have the main agent generate prompts for its own subagents [1]. Practical flow: define the top-level goal, ask for decomposition, then have the parent draft the child prompts before launch [1].

- **Front-load clarifications before you fan out.** Tell the agent to ask every ambiguity-filling question up front, then give it all required context so the run does not stop every few minutes for missing details [1]. This pairs directly with the small-context rule behind fan-out [1].
- **Require self-tests, then manage the fleet instead of one chat.** Have the agent generate its own integration sanity tests as part of the run [1]. In the same workflow, the human role shifts toward supervising many active agents rather than micromanaging one session [1].

## WHAT SHIPPED

- **Artifact worth studying:** Armin Ronacher surfaced a 280kLOC AI-generated pull request against WebKit and said it is a reminder that “loops are coming for core infrastructure” [2, 3].
- **Adoption signal:** the fan-out pattern above is already being used inside Cognition’s model research team to spin up 100 Devins on eval logs, and by product teams to run 5 child Devins against 5 alternative implementations of the same idea [1].

## GO DEEPER

- **Study the original workflow writeup:** imjaredz on Devin fan-out. It is a compact but unusually actionable thread on subagent orchestration: decomposition, child prompt generation, up-front clarification, parallel VMs, and agent-written sanity tests [1].
- **Read the maintainer reaction alongside the code:** Armin Ronacher’s note + WebKit PR #249. Useful if you are thinking about how established projects review and absorb very large AI-generated changes [2, 3].

“Seeing a 280kLOC AI generated pull request against WebKit is a good reminder that loops are coming for core infrastructure. It’s both exciting and confusing. I wouldn’t know how to run an established project and make that change.” [2]

*Editorial take: the edge is moving from better one-shot prompts to better decomposition—small contexts, parallel workers, and review processes that can handle much larger AI-generated diffs [1, 2].*

---

## Sources

1. X post by @imjaredz
2. X post by @mitsuhiko
3. X post by @mitsuhiko