Claude Moves to the Desktop as T3 Code, Cursor, and LangSmith Sharpen the Loop

Coding Agents Alpha Tracker

2026-03-24

Claude Moves to the Desktop as T3 Code, Cursor, and LangSmith Sharpen the Loop

By Coding Agents Alpha Tracker • March 24, 2026

Anthropic's Claude computer-use preview is the headline, but the sharper practitioner signal is the support stack around it: official CLI-based clients, faster repo search, and webhook-driven handoff for long-running agents. This brief also covers CodexBar 0.19.0, OpenClaw's latest beta, and the concrete workflows worth copying.

TOP SIGNAL

Anthropic pushed Claude past the repo window: the official Claude account says the new macOS research preview can open apps, navigate browsers, and fill spreadsheets in Claude Cowork and Claude Code, while Boris Cherny said Anthropic Labs is releasing full computer use in Cowork and Dispatch [1, 2].

Elsewhere, teams attacked the adjacent bottlenecks: T3 Code used the official Claude CLI, community contributors added browser control to its open-source UI, Cursor cut search latency across huge codebases, and LangSmith showed a webhook flow for long-running agents [3, 4, 5, 6].

> "The future where I never have to open up my laptop to get work done is becoming real very fast" [7]

TOOLS & MODELS

- **Claude computer use (research preview)** — Claude can now use your Mac to open apps, drive the browser, and fill spreadsheets. Officially this is a research preview in Claude Cowork and Claude Code on macOS; Boris Cherny said the release marks full computer use in Cowork and Dispatch,

and noted the early Sonnet 3.6 prototypes were clunky but already showed the use cases [1, 2].

- **T3 Code + Claude Code subscriptions** — If Claude Code is already installed locally, Theo says you can just run `npx t3` or use the T3 Code app; it talks to the local Claude Code CLI through Anthropic's Agent SDK, with no extra auth screen or API-key setup inside T3 Code [3]. Theo contrasts that with OpenCode's dropped Claude Max plugin, which he says relied on its own harness, custom auth flow, and faked headers [3]. He also calls out the economics: the Claude Code subscription is $200/month for up to $5,000 of compute [3].
- **Cursor Instant Grep** — Cursor says it can search millions of files and return results in milliseconds, which directly speeds up agent task completion. They also published a build writeup covering the algorithms and tradeoffs; Jediah Katz called it singular technical work and said this is why alternatives feel slow [5, 8]. Writeup: cursor.com/blog/fast-regex-search [9].
- **CodexBar 0.19.0** — New release adds Alibaba Coding Plan support, subscription history charts, Cursor Total/Auto/API dashboard alignment, Codex code-review reset times, and a broader Claude stability/refactor pass [10]. Release notes: v0.19.0 [10].

WORKFLOWS & TRICKS

- **Async completion alerts for long-running agents** — Hari's LangGraph/LangSmith flow is clean and reusable:
  1. Clone the Deep Research example from LangChain's Deep Agents repo [11].
  2. Create `webhook.py` with a FastAPI route that receives the LangSmith payload, reads `payload.values.messages[-1].content`, and POSTs that final AI message to a Slack webhook [11].
  3. Register the FastAPI app under the HTTP app field in `langgraph.json`, then run `langgraph dev` [11].
  4. Create a background run with your thread ID, assistant ID set to `research`, an input message, and the webhook URL; the result is a Slack summary plus the full report in LangSmith tracing [11]. Timeless pattern: don't poll long jobs—ship a webhook and move on [11]. Docs: LangSmith webhooks [6].
- **Route models by task, not by brand loyalty** — Theo says he uses `54` for most coding, then opens a *new thread* and switches to Claude for UI passes, quick tidy-ups, and small changes [3]. The constraint matters: once you pick Claude Code for a thread, he says you can't switch harnesses mid-thread because the thread state, compaction, and related data are tied to that thread in the cloud [3]. Practical takeaway: treat thread boundaries as routing boundaries.
- **Use Codex review as triage, not final judgment** — Peter Steinberger's PR loop is blunt: let Codex find issues, ask whether the issue

is actually clear, ask whether the proposed fix is the best possible one, then continue the tradeoff discussion and usually rewrite the PR [12]. His warning is the timeless part: overly local fixes make the codebase unmaintainable [12].

PEOPLE TO WATCH

- **Boris Cherny** — high signal because he is speaking from the Anthropic Labs shipping team. He says that team shipped MCP, Skills, Claude Desktop, and Claude Code, and is now rolling out full computer use [2].
- **Theo** — worth tracking because he is both shipping T3 Code and publishing the integration details: official CLI vs custom harnesses, subscription economics, and how he routes models across threads in daily use [3].
- **Peter Steinberger** — useful today for three separate practitioner signals: CodexBar 0.19.0, a concrete Codex PR-review loop, and OpenClaw plugin/release activity [10, 12, 13, 14].
- **Jediah Katz** — short post, strong signal from someone building Cursor's agent: Instant Grep is why other tools feel slow [8].
- **Hari from LangChain** — useful if you care about deployment mechanics, not just model chatter. Today's video walks through a full webhook-driven completion flow end to end [11].

WATCH & LISTEN

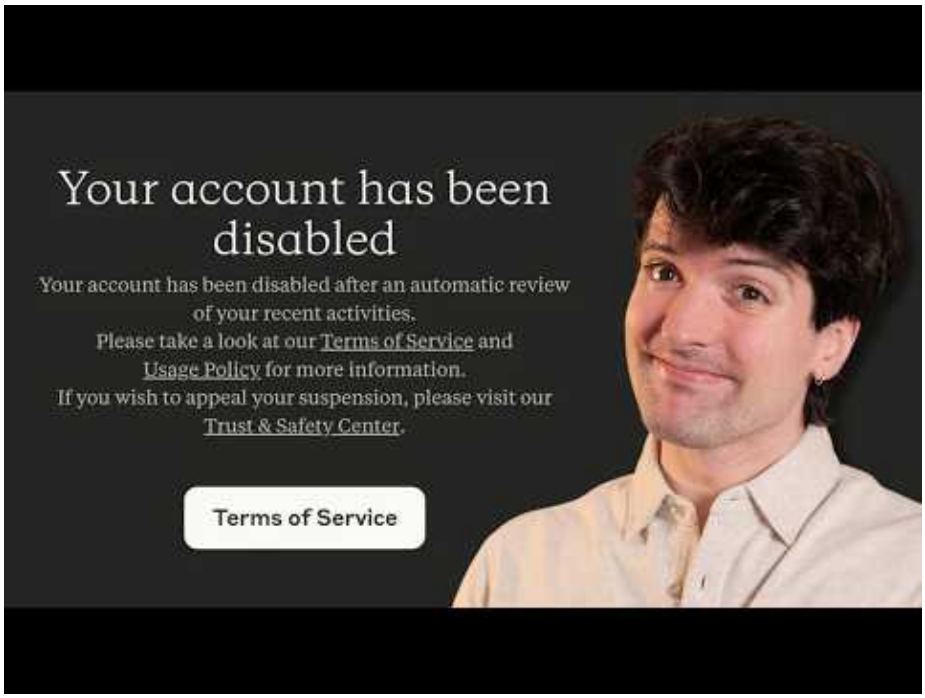
- **2:00-4:29 — Build the Slack webhook handler.** Hari shows the exact FastAPI route, the payload shape, and the one field that matters most: the final message at `values.messages[-1].content` [11].

Webhooks for LangSmith Deployment: Notify Slack When Your Agent Run Finishes (1:59)

- **5:24-7:11 — Kick off a background run with a webhook URL.** This is the concrete API/docs walkthrough: create a thread, call background run creation, pass the webhook endpoint, and wait for the Slack ping instead of babysitting the job [11].

Webhooks for LangSmith Deployment: Notify Slack When Your Agent Run Finishes (5:23)

- **12:47-13:15 — Why T3 Code built a harness abstraction.** Theo explains the real integration problem: every CLI exposes events differently, so supporting multiple providers means normalizing their weirdness instead of pretending the harness layer doesn't matter [3].



I need you guys to trust me on this (sorry Anthropic) (12:47)

PROJECTS & REPOS

- **T3 Code** — The open-source UI keeps picking up contributions: a community contributor added browser integration, terminal support is next, and the main app now supports Claude Code subscriptions through the local CLI path [4, 15, 3].
- **OpenClaw** — New beta `v2026.3.22-beta.1` is out. Separately, Harold connected Codex App Server to OpenClaw via plugins, and steipete highlighted that as a plugins story worth watching [14, 16, 13]. Release notes: v2026.3.22-beta.1 [14].
- **Deep Agents repo** — LangChain's webhook demo uses the Deep Research example from this repo; if you want to copy the same background-run pattern, it's the repo Hari recommends cloning locally [11].

Editorial take: today's edge wasn't a benchmark bump; it was better plumbing—desktop control, faster search, official harnesses, and async completion hooks that make agents usable in real workflows. [1, 5, 3, 11]

Sources

1. X post by @claudeai

2. X post by @bcherny
3. I need you guys to trust me on this (sorry Anthropic)
4. X post by @LLMJunky
5. X post by @cursor_ai
6. X post by @LangChain
7. X post by @alexalbert___
8. X post by @jediahkatz
9. X post by @cursor_ai
10. X post by @steipete
11. Webhooks for LangSmith Deployment: Notify Slack When Your Agent Run Finishes
12. X post by @steipete
13. X post by @steipete
14. X post by @steipete
15. X post by @theo
16. X post by @huntharo