OpenClaw After Steinberger — What the OpenAI Move Means for Your Setup

February 27, 2026 · by Mark Bartlett

Download this guide as PDF

> **Quick Answer:** OpenClaw is not dead. Three releases shipped in the week after Steinberger joined OpenAI (2026.2.22 through 2026.2.24), with real security fixes and new features. The project is moving to an independent foundation, MIT license stays, and 859 contributors are still pushing code. The npm package hit 1.27 million weekly downloads on Feb 25. Update to 2026.2.24 or later for the security hardening, review your channel permissions, and keep building. The panic is louder than the problem.

📚 **Related:** OpenClaw Security Guide · OpenClaw Setup Guide · Best OpenClaw Alternatives · How OpenClaw Works · ClawHub Security Alert

Two weeks ago, OpenClaw's creator Peter Steinberger joined OpenAI. Since then, the project has shipped three releases, Elon Musk posted a monkey-with-a-rifle meme about it, Meta's AI safety director had her inbox deleted by her own OpenClaw agent, Baby Keem asked Twitter how to fix internal reasoning leaking, and Perplexity launched a competitor.

If you saw any of that and wondered whether to uninstall OpenClaw, keep reading. The short version: no.

We covered the initial acqui-hire announcement when it happened on February 15. This article is the practical follow-up. What has actually changed in the codebase, what the foundation transition means for your setup, and which of the noise floating around is worth paying attention to.

What actually happened

Peter Steinberger announced on February 14 that he's joining OpenAI to work on personal agent technology. In his blog post, he wrote that OpenAI is "the fastest way to bring this to everyone" and described himself as "a builder at heart" rather than a company founder. Sam Altman called personal agents "core to our product offerings."

OpenClaw itself was not acquired. The project moves to an independent open-source foundation. The MIT license stays. OpenAI committed to sponsoring the project and giving

Steinberger time to contribute. Steinberger framed the foundation as “a place for thinkers, hackers and people that want a way to own their data.”

The governance model is shifting from BDFL (Steinberger making final calls) to a maintainer council. If you’ve watched Linux or Kubernetes go through this same transition, you know how it ends. The project gets more stable, not less.

Repository numbers as of February 26: 230,815 stars, 44,259 forks, 859 contributors, 1.27 million weekly npm downloads.

Three releases shipped after Steinberger left

This is the thing that matters most if you’re deciding whether OpenClaw has a future. Talk is cheap. Releases aren’t.

2026.2.22

Shipped February 23. Session isolation improvements, multi-account cron routing fixes, typing indicator stability across Discord and Telegram.

2026.2.23

Shipped February 24. This one is the security release you should care about:

- **SSRF protection**: Browser fetch now defaults to trusted-network mode, blocking server-side request forgery attacks
- **Sandbox hardening**: Docker container namespace-join mode blocked for sandbox containers
- **HTTP security headers**: Strict-Transport-Security for direct HTTPS deployments
- **Credential redaction**: Prevents prompt injection from extracting stored credentials
- **Session cleanup**: Disk budget controls to prevent transcript storage from growing indefinitely

Also added Claude Opus 4.6 support and Kilo Gateway integration.

2026.2.24

Shipped February 25. Auto-reply abort shortcuts expanded with multilingual support (Spanish, French, Mandarin, Hindi, Arabic, Japanese, German, Portuguese, Russian). Heartbeat now blocks DM delivery by default. Multi-user heuristic detection for security audits. Session isolation fixes for cross-channel replies.

Two more releases (2026.2.25 and 2026.2.26) shipped after that, covering Android streaming improvements, external secrets management, and agent routing CLI commands.

Five releases in five days. Whatever else is happening around the project, the code is moving.

The Musk tweet and the inbox incident

On February 23, Meta's AI safety director Summer Yue shared a story on X: she gave OpenClaw access to her work inbox, told it to "suggest what you would archive or delete," and watched it speedrun deleting everything. She had to physically sprint to her Mac mini to kill the process.

Elon Musk responded with a meme of a monkey being handed a rifle, captioned: "People giving OpenClaw root access to their entire life." The post hit 48,000 engagement.

The security concern here is legitimate. It is also not new.

OpenClaw connects to WhatsApp, Telegram, Discord, Gmail, your file system. It executes actions autonomously. If you give it access to your inbox and tell it to do something ambiguous, it will interpret that ambiguity and act on it. That's the entire point of an autonomous agent, and it's also the entire risk.

The 2026.2.23 and 2026.2.24 releases address exactly this. SSRF guards now block unauthorized network requests. Docker sandbox isolation got tighter. And the heartbeat change in 2026.2.24 means agents can no longer send unsolicited DMs by default — you have to explicitly opt back in.

None of this makes OpenClaw "safe" in the way that a calculator is safe. You are running an autonomous agent that takes actions on your behalf. If you give it access to your email, it can delete your email. The fix is the same as it's always been: don't give agents permissions you aren't comfortable with them using. Our OpenClaw security guide covers this in detail, and the ClawHub security alert covers the community marketplace risks.

Musk's tweet was dunking on a rival's hire, not a security analysis. The underlying concern is real, but the people working on OpenClaw's codebase are shipping actual fixes faster than the people tweeting about it.

Baby Keem and mainstream penetration

On February 25, rapper Baby Keem posted: “how do u fix openclaw internal reasoning leaking.” The tweet hit 1.2 million views and 16,000 likes. Theo (t3.gg) quote-tweeted it with: “Baby keem is using openclaw and you’re still writing code by hand.”

When musicians are debugging agent behavior on Twitter, you’re past the early adopter phase. The community is going to get noisier, the bug reports are going to get less technical, and the project needs governance that can handle that scale. A foundation is the right structure for that.

Perplexity entering the ring

Perplexity launched “Computer” in the last week of February, a managed AI agent platform that coordinates 19 models on the backend. It’s positioned as OpenClaw for people who don’t want to self-host. The pricing: $200/month on the Perplexity Max plan.

The timing is strategic. Steinberger leaves, uncertainty spikes, Perplexity launches a competitor. That’s normal market behavior.

	OpenClaw	**Perplexity Computer**
Runs locally	Yes	No (cloud only)
Price	Free (open source)	$200/month
Data ownership	You own everything	Perplexity’s servers
Model choice	Any (Ollama, OpenAI, Claude, etc.)	19 models, Perplexity chooses
Setup	Manual (or OpenClawd managed)	One-click
Risk profile	You control the boundaries	Perplexity controls the sandbox

If you’re reading InsiderLLM, you probably care about running things on your own hardware. Perplexity Computer solves a real problem for people who want agents without managing infrastructure. It’s not a replacement for self-hosted OpenClaw, and the $200/month price point makes it a non-starter for hobbyists.

OpenClawd, a third-party managed platform, offers hosted OpenClaw for less than Perplexity charges. They launched on February 9 and updated their platform on February 20, explicitly timing the release to the Steinberger news. It’s the same OpenClaw codebase with security

defaults applied and no terminal required. OpenClawd is not affiliated with OpenAI or the OpenClaw foundation.

More competition means more options. OpenClaw's advantage hasn't changed: your hardware, your data, your rules.

What the foundation transition means for your setup

Practically: almost nothing changes right now.

License

MIT. No change planned. Steinberger committed to this in his blog post and OpenAI confirmed it publicly.

Release cadence

If anything, slightly faster. The project shipped five releases in the five days after Steinberger's departure. The maintainer council hasn't formalized yet, but 859 contributors don't need one person's approval to merge code.

Breaking changes

The 2026.2.23 and 2026.2.24 releases do include breaking changes:

- **SSRF policy default**: Browser fetch now defaults to trusted-network mode. If you're running OpenClaw on a private network with local services, run `openclaw doctor --fix` after updating.
- **Heartbeat DM blocking**: Agents can no longer send direct messages by default. If your setup relies on the agent proactively messaging you, add `directPolicy: allow` to your heartbeat config.
- **Docker sandbox isolation**: Namespace-join mode is blocked. If you were using it for performance, you'll need to switch to standard isolation.

These are security-motivated changes. They may break workflows that depended on permissive defaults. The fix for each is documented in the changelog.

Commercial ecosystem

Third-party businesses are building on OpenClaw: OpenClawd (managed hosting), MyClaw.ai (one-click deployment), and various integration providers. A healthy commercial ecosystem around an open-source project usually means the project survives. Companies don't build products on top of dead code.

What you should do right now

Update to 2026.2.24 or later

The security hardening in these releases is real. SSRF protection, credential redaction, and sandbox fixes address the class of vulnerabilities that the Summer Yue incident highlighted. Don't skip this.

```
# Check your current version
openclaw --version

# Update via npm
npm update -g openclaw

# Or via Homebrew
brew upgrade openclaw
```

Review your channel permissions

Go through each connected channel (Telegram, Discord, WhatsApp, Gmail) and verify the agent only has access to what you actually need. If it has your email and you don't use email features, revoke that.

Back up your config and sessions

```
# Back up your OpenClaw directory
cp -r ~/.openclaw ~/.openclaw-backup-$(date +%Y%m%d)
```

Do this before major updates. Session data, custom skills, and channel configs are all in `~/.openclaw/`.

Run the diagnostic

```
openclaw doctor --fix
```

This catches misconfigurations, especially after the SSRF policy change in 2026.2.23. Run it after every major update.

If you're new: this is actually a fine time to start

The hype means better documentation, more community guides, and faster bug fixes. The setup guide covers installation from scratch. Security is better now than it was a month ago because the scrutiny forced real improvements.

Perspective from someone who's watched this before

Open-source projects survive founder departures all the time. Guido van Rossum stepped back from Python in 2018 and the language kept growing. Torvalds took a break from Linux the same year and kernel development didn't slow down. The pattern that kills projects is a founder leaving behind a small contributor base, no governance, and no commercial ecosystem. OpenClaw has none of those problems.

What could go wrong: the foundation governance takes too long to formalize, key maintainers burn out, or OpenAI's sponsorship creates conflicts of interest. These are risks worth monitoring, not reasons to jump ship.

I would be more worried if the releases had stopped. They haven't. Five releases in five days, with actual security improvements. That's a project with momentum.

Bottom line

The project is moving faster than the discourse about it. Five releases shipped in the first week without Steinberger. The MIT license stays. The foundation transition follows the same playbook that worked for Linux and Kubernetes.

The security concerns Musk amplified are legitimate. They're also being fixed in actual code, which puts the OpenClaw maintainers ahead of most people tweeting about it. Update to 2026.2.24 or later, review your permissions, and keep building.

If 859 contributors and 1.27 million weekly downloads can't sustain a project, nothing can. OpenClaw is fine.