

The Grok Bot Field Guide

Verdict first. Risk second. Hype last. Everything else buries this in the wrong order.

Compiled from official documentation, hands-on setups, independent reviews, and business use-case research.
August 2026

1. The One-Paragraph Verdict	3
2. What It Actually Is	3
3. What It Actually Costs	4
4. What Nobody Tells You First	5
5. Real Setup, Step by Step	6
6. Every Gotcha, Compiled	7
7. What To Actually Point It At	8
8. How It Actually Compares	10
9. When Something Breaks	11
10. The Verdict, Again	11

1. The One-Paragraph Verdict

Grok Bot is a real, working product for delegating draft-first work — inbox triage, research, content, internal reporting — to a team of named AI agents that each get their own cloud computer and sign into your tools like a person would. It costs \$200-\$300/month with no spend cap on overage yet. **It's ready now for internal, reversible work. It is not ready for anything that reaches a customer without a human checking first** — the controls that would make that safe (rehearsal, defined scope, an audit record) aren't fully shipped. If your work is mostly drafts and research, this earns its cost fast. If you're hoping to hand it your support queue unsupervised on day one, wait.

2. What It Actually Is

Grok Bot launched in beta on August 11, 2026 — a joint SpaceXAI and Cursor product, built after SpaceX's acquisitions of both xAI and Cursor closed earlier in 2026. It runs on the Grok 4.6 model family, tuned specifically for long-running agentic work rather than conversation.

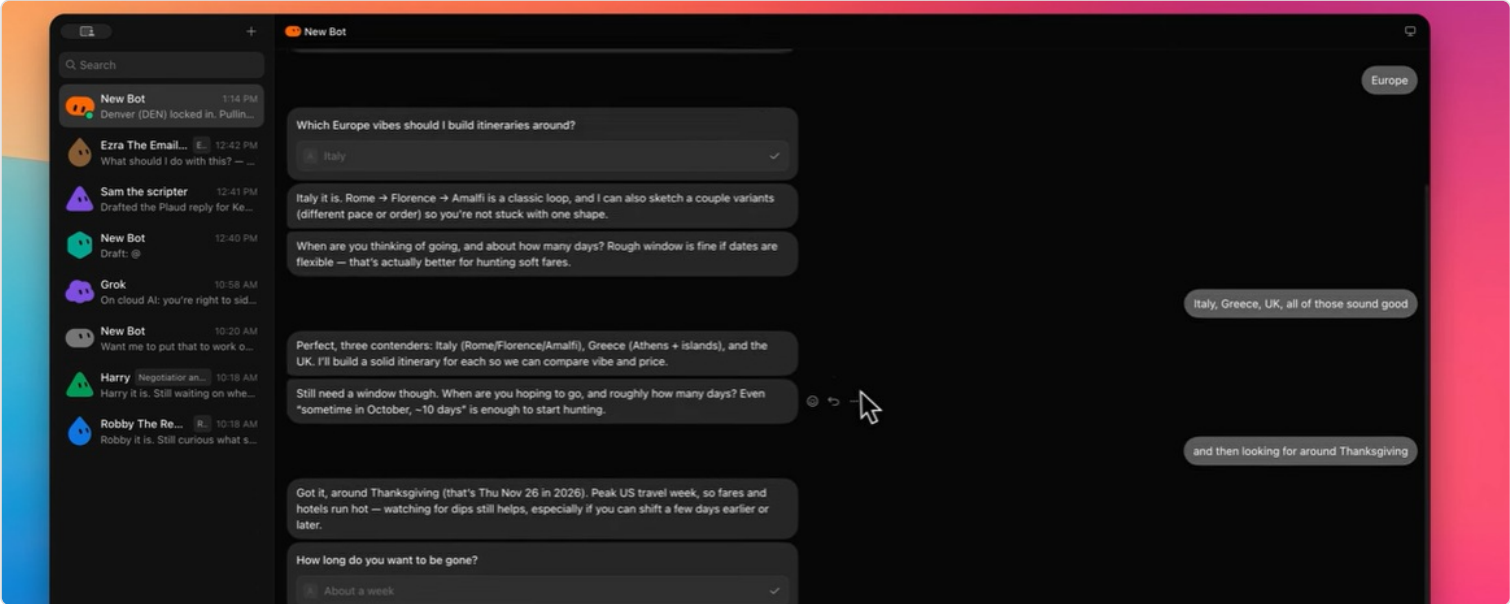
Three mechanisms explain almost everything it can and can't do:

It has its own computer

Each Bot runs on a persistent cloud VM with a browser, filesystem, and terminal — not a process on your laptop. Work continues after you close your laptop. Multiple Bots you create share one user-scoped computer and can run in parallel.

It signs in like a person, not an API

For sites without a plugin, the Bot opens a browser on its own cloud computer and clicks through the interface the way you would — which means it also works on old software that never had an API. You sign in yourself when it hits a password or 2FA prompt; the Bot never sees your password.



A real Grok Bot conversation — the bot asks one clarifying question at a time rather than guessing, with quick-reply suggestions under each message.

It learns by watching

Ask a Bot to follow along as you complete a workflow once. It saves the sequence as a routine and runs it on its own next time — including a genuinely notable self-correction behavior: in one documented setup, an agent reviewed its own output after a run, flagged an ambiguity it had noticed on its own, and adjusted its own instructions to fix it without being asked.

"An agent that gets a task right forty-nine times will do something odd on the fiftieth, with full confidence and no warning."

3. What It Actually Costs

Grok Bot has no standalone price. It's bundled into exactly three paid tiers — and two cheaper, similarly-named Grok plans do not include it at all, a detail almost every launch article buried in one sentence or skipped entirely.

SuperGrok

\$30 USD/month

For the most demanding coding work

Upgrade to SuperGrok

By subscribing, you agree to our Terms and Privacy Policy.

</>

 Access to Grok Build

📄

 Flicker-free CLI

⌘

 Parallel subagents and worktrees

💡

 Smarter planning and self-review

🔗

 Extend with MCP, skills, plugins, and hooks

SuperGrok Plus

\$100 USD/month

Significantly higher usage

Upgrade to SuperGrok Plus

By subscribing, you agree to our Terms and Privacy Policy.

🔄

 Everything in SuperGrok

🔗

 Create apps with a single prompt
With one-tap deploy to the web

📺

 Create 1080p videos

🗣️

 Significantly higher usage across Chat, Imagine, Voice & Build

⚡

 Lightning-fast replies

SuperGrok Heavy

\$300 USD/month

World's most powerful AI

Upgrade to SuperGrok Heavy

By subscribing, you agree to our Terms and Privacy Policy.

👤

 Exclusive access to Grok Bot

🔄

 Everything in SuperGrok Plus

🚀

 Highest usage at the fastest speed

🔍

 Solve extremely hard problems

🧠

 Most powerful intelligence
Larger team of agents collaborating to get you

xAI's own pricing page. Grok Bot access is listed under SuperGrok Heavy only — labeled "Exclusive access to Grok Bot." Not SuperGrok, not SuperGrok Plus.

Plan	Price	Includes Grok Bot?
Free	\$0/month	No
SuperGrok	\$30/month	No — Grok Build access only
SuperGrok Plus	\$100/month	No — 1080p video, higher usage
SuperGrok Heavy	\$300/month	Yes — exclusive tier
Cursor Ultra	\$200/month	Yes — cheapest solo entry point
Cursor Teams Premium	\$120/seat/month	Yes — cheaper per seat than solo

The bigger risk isn't the sticker price. Each plan includes a weekly usage allowance, but there's no Grok Bot-specific spend cap published yet. Go over the allowance and you're billed on demand from raw model and token cost — around \$2 per million input tokens and \$6 per million output tokens on the underlying model, doubling above 200k prompt tokens. There's no model picker to steer heavier tasks toward something cheaper. Pilot this on a low-stakes week before assuming your monthly number is fixed.

One real arithmetic check worth doing before you subscribe: a daily task that saves an hour a day covers a \$200 monthly plan within the first week. That's the actual test — not whether the tool is impressive, but whether you can already estimate the hours it would save, because that makes the payback checkable instead of hopeful.

4. What Nobody Tells You First

Every casual setup guide leads with the feature list. This is the section that actually decides whether your first month goes well.

Every bot you create shares one computer

According to xAI's own documentation, isolation is at the **user-account level**, not per-bot. Files, browser sessions, and command-line credentials are available across your entire roster. Each bot gets its own screen — that's a separate work surface, not a separate security boundary.

xAI's own wording, worth repeating exactly: "Do not use separate Bots as a security boundary." And: "Deleting a Bot does not remove shared-computer files or browser sessions." Deleting a bot is not the same as revoking what it had access to.

An agent with a login is an employee with a login

Four rules the independent field research is explicit about, and almost none of the public excitement mentions:

- **Give it its own account.** If it logs in as your ops manager, your audit trail says your ops manager did it. For anything a regulator might ask about later, that's a real problem.
- **Check your software licenses.** Software priced per seat is priced for people. An agent using your team's access may not be permitted under the agreement you signed — nobody else covering Grok Bot flagged this, and it's a genuine compliance question.
- **Start narrow.** Read widely, write rarely, require human approval for anything that spends money or leaves the company. Widen after a month of watching, not before.
- **Sort work by the cost of a mistake.** Keep a person between the Bot and anything expensive or customer-facing.

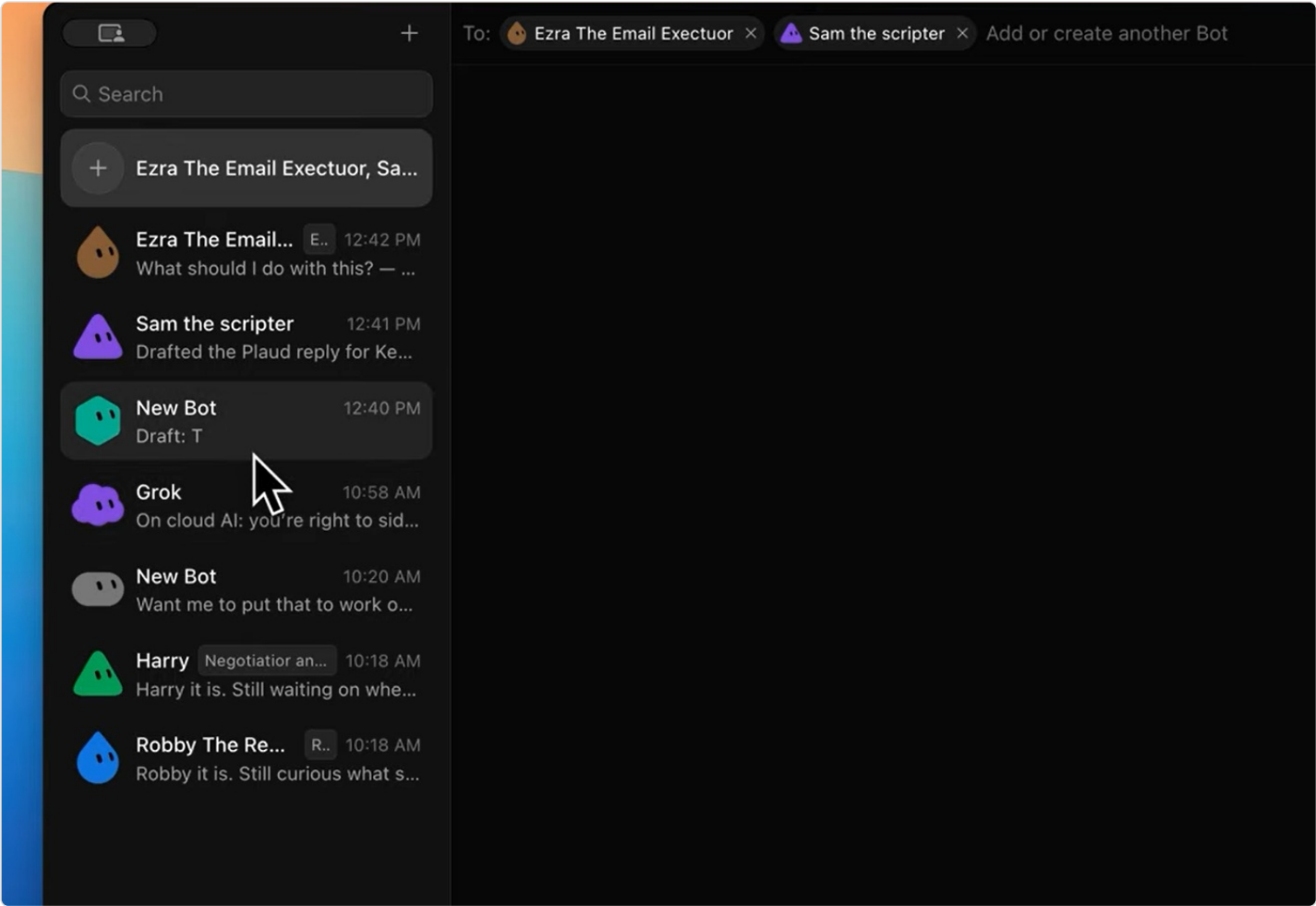
It cannot reach anything on your own machine

Confirmed directly by Cursor staff: "Grok Bot does not attach MCP servers that run on your own machine, whether that's stdio or something listening on localhost." If a workflow genuinely needs a local process, that's Grok Build — a separate product — not Grok Bot.

What independent reviewers actually recommend

Use case	Verdict
Drafts, research, internal work	Ready now — this is where it's strongest
Anything that reaches a customer directly	Not yet — rehearsal, defined scope, and an audit record aren't fully shipped

5. Real Setup, Step by Step



A real bot roster — named and colored by role ("Ezra The Email Executor," "Sam the scripter"). Naming by job is not cosmetic; it helps the bot infer intent and helps other bots route work correctly.

- 1 **Download and sign in.** Grok Bot is a desktop and mobile app. Sign in with the Cursor account that should own your plan and usage.
- 2 **Create your first bot — make it the Chief of Staff.** Every account that started here had a smoother setup than accounts that jumped straight to five specialists.
- 3 **Connect your apps through Plugins.** Only what you're about to actually use. Connections are shared account-wide once authorized — connect Gmail once, every bot you create afterward can use it.
- 4 **Hand the Chief of Staff the real job: designing your team.** This is the step almost everyone skips.
- 5 **Review its recommendations before creating anything.**
- 6 **Only add a new bot when a real task is getting in your way.** Not one per subtask.
- 7 **Organize by client or project with Sidebar Sections.** Syncs between desktop and iOS. Requires app v1.2.0+.

The exact prompts

OPTION A — QUICK-START (USED LIVE, ON CAMERA, BY A DEVELOPER SETTING THIS UP FOR THE FIRST TIME)

```
take a look at me and what i do. find all info you can on me. take a look at the bots i have created. these are your team. see if we need to change anything or add more bots. what is best way of managing these bot teams. what else will make me super productive
```

OPTION B — CONNECTIONS-FIRST (BEST ONCE YOU'VE ALREADY CONNECTED SEVERAL APPS)

```
Review my connections, build a profile of my workstreams and projects, then suggest the three most useful agents, automations, and any plugins I should add.
```

OPTION C — SUPERVISED (RECOMMENDED FOR ANYTHING CLIENT-FACING)

```
Set up a new bot for me in its own dedicated chat that acts as my chief of staff. Walk me through connecting the tools I use daily, then configure it to give me a morning summary covering important messages, meetings, commitments, and open work while flagging anything urgent. Ask me which people, projects, and urgency rules matter, run the first briefing with me watching, and require my approval before it sends messages, changes calendar items, or takes any external action; then save it for a daily run.
```

Two prompts worth stealing once you're set up

Turn a daily digest into a short morning podcast covering emails, calendar, and anything else I need to know for my day.

I have a meeting now, right? What's the link?

Real hard limits, so you don't design around a ceiling you can't see

Limit	Value
Bots per account	50
Routines per bot	50
Run records kept per routine	20 most recent

If you're planning a large fleet across multiple clients or projects, hit these numbers before you commit to an org structure — restructuring 40 bots later is a much bigger job than planning around the ceiling now.

Organize by client or project: Sidebar Sections

Group bots by project, client, or business without mixing them together. On iOS: swipe a bot row in the sidebar, tap **Move to**, choose **New Section** and name it. Repeat for other bots in the same group. Sections sync between iOS and desktop. Deleting a section moves its bots to Unassigned rather than deleting them. Requires app v1.2.0 or later.

Track what your bots are actually doing: the project-manager trick

Once you're running more than two or three bots, keeping track of what each one has done stops being obvious from the chat list alone. One workable fix, used by an agency running several bots at once: connect a task-management tool (ClickUp is the one most commonly used for this) and have your Chief of Staff log every agent task there as it happens. You get one board showing what every bot is working on, what's done, and what's stuck — instead of scrolling back through separate chat threads to reconstruct it.

6. Every Gotcha, Compiled

Scattered across a dozen separate blog posts elsewhere. Compiled here once.

- **GitHub can ask for a token even when it's already connected in Cursor.** Skip the second connector and sign in directly on the bot's own computer instead.
- **X requires a bearer token** from the developer portal. Budget ten minutes the first time, or skip X until later.
- **LinkedIn posting can take a few minutes to go live.** Normal, not a failure.
- **LinkedIn won't let you attach an image to a post after it's live.** Add a screenshot as the first comment instead.
- **Team/group bot chats currently can't save routines.** You have to go into each bot's individual chat to create one.
- **No local MCP, ever.** Grok Bot cannot reach localhost or a private IP on any machine. Use Grok Build if a workflow needs that.
- **MCP custom connectors have no static OAuth client ID field** — blocks some enterprise MCP servers today.
- **For apps with no native plugin, use Composio.** Sign up separately, connect the app there instead of through Grok Bot's own Plugins list, and your bot can reach it through that bridge automatically on the next request — this is exactly how one demonstrated setup pulled YouTube analytics with no native YouTube connector available.
- **Have bots save or sync work once at the end of the day** rather than continuously — frequent connector updates burn more usage for no real benefit.

7. What To Actually Point It At

Most coverage picks the most impressive-looking use case. That's the wrong filter. Use these four tests instead.

- 1. It crosses two systems.** Work stuck between an inbox and a CRM, or a spreadsheet and an accounting tool, is where agents earn their keep. Single-system work usually has a cheaper fix.
- 2. Someone does it daily.** A monthly task saves an hour a month. A daily one changes someone's week — and gives you fifty times the evidence about whether the agent is actually reliable.
- 3. A mistake is visible and cheap.** A wrong draft a person reviews costs nothing. A wrong payment costs real money and trust. Start where errors surface immediately.
- 4. The rule can be written down.** If you could hand the instructions to a new employee on paper, an agent can attempt it. If the rule lives only in one person's judgment, start elsewhere.

Worked example: a logistics company answering 30 "where is my shipment" emails a day. Crosses two systems (inbox, tracking). Happens daily. Mistakes are visible and cheap (a wrong reply gets caught by the customer immediately). The rule is writable (look up reference, report status, escalate anything unusual). Passes all four — which is why variations of this exact task keep showing up wherever these agents are actually working.

30 real, observed use cases — not hypothetical marketing

Every one below was publicly documented by an early beta user who actually watched it complete. Grouped by the work they replace.

Sales & Growth

- Book meetings by writing outbound emails in your own voice, then sending the notes for you
- Take a sales play that used to take 90 minutes and make it ~95% automated
- Rebuild most of a custom CRM with a small Bot crew in about a day and a half
- Screen 1,000+ event applicants against a target profile and batch-approve the fits
- Review 200 job applications into strong, mid, and reject in one pass
- Answer a Salesforce end-of-quarter question from your phone in ~10 seconds
- Search company chat for who asked about a feature, in 15 seconds instead of an hour
- Turn a week of sales-call recordings into a win/loss memo with the phrases that closed
- Forecast account health and catch churn early when usage is quietly dropping
- Watch a competitor's pricing page and message you the moment it changes
- Work around a missing permission in a sales tool by driving its screen directly
- Record a workflow once, have the Bot turn it into an automation

Needs: inbox and CRM access, a clear written definition of a good outcome. The screening/reviewing tasks are easiest to start with — a person checks the output anyway.

Engineering & Systems

- Log into a monitoring tool and read performance data no API exposes
- Watch a database migration step by step, checking logs at each stage
- Run a test suite across AI models and compare quality against cost
- Flag conflicting code changes across a busy repository
- Scrape a topic from social media into a spreadsheet with exact columns
- Audit your whole Bot fleet for wasted spend and shut down the wasteful ones

Needs: read access only, almost nothing else. Safest place to start technically — nearly everything here is observation, not action.

Internal Operations

- Deliver one morning digest across email, chat, calendar, and call notes
- Find 90 minutes of focus time this week and decline the junk meetings
- Triage email twice a day
- Filter a noisy chat channel down to only what needs action
- Rebuild a messy shared drive into a wiki with named owners
- Write a first-week plan for a new hire from the last few people's calendars
- Propose two automations by inspecting your tools' history, unprompted

Needs: broad read access, very little write access — low risk. The real catch: value depends on the person actually changing their habits. A digest nobody reads automates nothing. Start with one person who genuinely wants it.

Finance & Support

- Track invoice actuals for large expenses with a bookkeeping routine
- File expenses from a photo dump of receipts, coded to the right category
- Run an hourly pass on the support queue, escalating only what needs a person
- Turn a week of support tickets into changelog notes and help-center articles

- Monitor incident-related ticket volume in the background

Needs: a written rule and a human checkpoint. Keep a person approving anything customer-facing for the first month, then narrow the review to exceptions.

Does this work on old software with no API?

Usually — and it's the most commercially important fact in this section. Because a Bot signs in rather than calling an API, it reaches systems that never had one: the aging ops tool, the supplier portal nobody's touched in years, the accounting system that was modern in 2016. That reverses the usual order of operations — you can now test automation against the system you already have and decide about modernizing it with evidence instead of a guess. The honest caveat: driving a screen is slower than an API and breaks when the interface changes. Treat it as a cheap test, not permanent architecture.

8. How It Actually Compares

	Grok Bot	Claude Cowork	ChatGPT
Maker	SpaceXAI + Cursor	Anthropic	OpenAI
Remote MCP	Yes, cloud VM only	Yes, Free through Enterprise	Yes, write actions need Business+
Local stdio MCP	Not in Grok Bot itself (Grok Build supports it)	Yes, via Claude Desktop / Claude Code	Not in ChatGPT itself (Codex supports it)
Pricing floor for write-capable use	\$200-300/month	Free tier, one connector	\$20/user/month, two-seat minimum

	Grok Bot	Self-hosted (OpenClaw / Hermes)
Setup	Minutes, guided	Hours, technical
Who hosts it	SpaceXAI — you don't own the harness	You do, full control
Cost model	Bundled subscription + uncapped overage	Infrastructure + API costs, most predictable at scale
What you babysit	Just the approval queue	The gateway, the config, the model

A fourth model worth knowing: Buzz

Buzz isn't a teammate with its own computer — it's closer to a shared room. You host a relay (or pick a public one), spin up as many agents as you want inside it, and it functions as a human-plus-agent workspace with an audit trail, not a single named worker. The tradeoff is a real one: Buzz has a harder initial setup (saving a private key, which can be genuinely confusing for a non-technical user), but once past that, spinning up new agents inside it is fast — comparably fast to Grok Bot itself, according to hands-on comparisons. Worth evaluating if what you actually want is a shared space your whole team drops agents into, rather than individually managed bots each running their own machine.

Claude Cowork is positioned by independent reviewers as the closest direct analogue to Grok Bot — both market themselves as an AI coworker, built on opposite architectural bets (local + flexible vs. cloud + seamless). Worth evaluating both before committing a monthly budget to either.

9. When Something Breaks

Most issues clear with a full restart — fully quit the app (on Mac, Quit from the menu bar, not just closing the window), reopen, try again before anything else.

Symptom	Try first	If it persists
Computer reset or update stuck	Update the app, fully quit and reopen, give it time	Use Reset only as a last resort — unsynced work won't come back
A routine isn't running	Confirm it's enabled and the schedule is correct	If enabled but silent 24h+, report a bug with the agent name and routine details
Error or crash mid-run	Note the exact error text and request ID	If the same prompt fails again, report it with that info
Plugin won't authorize	Re-add it, finish the provider login in browser	If still disconnected, contact support with the plugin name and account email
App hangs on launch	Fully quit and relaunch	If it hangs every time, contact support with platform + version

10. The Verdict, Again

Grok Bot is a real product, not a demo — the cloud-computer architecture and the "learns by watching" mechanism are genuinely useful, and the 30 use cases in this guide are things people actually watched it do, not marketing copy. The risk isn't that it doesn't work. It's that the excitement around it is outrunning the governance conversation — shared credentials across your whole roster, no spend cap, license questions nobody's asking, an audit trail that can point at the wrong name. Use the Four Tests before you pick a task. Keep a person between it and anything expensive. Everything else in this guide exists to make that one paragraph actually actionable.

11. In a Hurry? One Prompt Does Most of This

If reading the full setup feels like too much right now, or you just want to see it work before committing — copy the whole block below into your first bot and let it walk you through the rest.

PASTE THIS WHOLE BLOCK AS ONE MESSAGE TO YOUR CHIEF OF STAFF BOT:

take a look at everything you can find about me and how I work, once you've connected to gmail and calendar (I'll sign in when it asks), build a full profile of my workstreams, then come back with exactly 3 recommended agents, what each one would own, and what it needs access to — ranked by which one saves me the most time first. don't create anything yet, just show me the plan.

This won't create anything automatically — on purpose. Signing into gmail/calendar happens in a real browser window it opens for you, and "don't create anything yet" is there so you see the actual plan before anything gets built. That pause is the whole point of the security section above — worth keeping even in the fast version.

Sources this guide draws from:

xAI official documentation and pricing pages (x.ai/bot, docs.x.ai) · Cursor's official Grok Bot help docs · posterly's footnoted Grok Bot platform guide · Rundown University setup guide · Debbie O'Brien's hands-on setup account (debbie.codes) · botdirectory.ai prompt library · eesel AI's independent review and pricing analysis · Latent Digital Solutions' 30 use-case business research · MacRumors, Interesting Engineering, and CellCog/MoClaw launch and pricing coverage · hands-on comparison walkthroughs covering Grok Bot, Buzz, OpenClaw, and Hermes side by side

This guide compiles and synthesizes publicly available documentation, hands-on accounts, and independent research. It is not affiliated with xAI, SpaceXAI, Cursor, or Anysphere.

Compiled by KYD — building ForgePilot.

Need more resources like this? I post new guides and tools personally — follow along on X: x.com/K_Y_D_A_A_L, or check everything already live at forgepilot.io/foundry