

ROHIT KUSHWAHA

AI Engineer · Generative UI, Agent Runtimes & Developer Tools
Ambarnath, India | rohitk290106@gmail.com | github.com/DevRohit06 | linkedin.com/in/rohitk06 | rohitk06.in

SUMMARY

I work on the unglamorous half of AI products: generative UI that renders what a model emits, tool calls that return the types they promised, and sandboxes that stop leaking under load. 294 of my pull requests are merged into PocketPaw, a self-hosted agent platform; the rest of my time goes into developer tools I ship myself.

EXPERIENCE

- Qbtrix** — Full-Stack Engineer Dec 2024 – Present
- Build **ripple**, Qbtrix's generative UI runtime: an agent emits a JSON spec of typed widget nodes instead of writing code, and the runtime renders it as a live canvas. Shipped widget waves 1–3 across a 150-widget catalog, from primitives up to composed enterprise layouts, with two-way binding and a **\$source** resolver so widgets read live data.
 - Made ripple's widget manifest the agent's single source of truth and wrote the validator that rejects a malformed spec before it reaches the renderer. That gate is what separates "the model produced some UI" from something you can put in front of a customer.
 - Authored **soul-protocol**, an open standard for portable AI identity and memory, so a user's persona and preferences survive a change of assistant instead of being locked to one vendor.
 - Own the frontend and the state layer under it across several Qbtrix products.

- Qbtrix** — Full-Stack Developer Intern Dec 2023 – Dec 2024
- Built the Interacly frontend and took it to production.

AI ENGINEERING WORK

- PocketPaw** — core contributor · self-hosted open-source agent platform, Python github.com/pocketpaw/pocketpaw
- Fixed a class of tool-calling bugs: the default backend reported every tool call with empty arguments, pydantic-ai dropped tool arguments before they reached the tool, and bridged tools received strings where they had declared ints.
 - Made the runtime send an explicit max_tokens, so a one-line reply stops being billed like an essay, and rewrote the tool descriptions in plain English, because a model picks the right tool from the description or not at all. Reworked the streaming surface to say what the agent is doing rather than name the function it called.
 - Queued sandboxed builds behind a gate that checks artifact bytes instead of trusting a sentinel, and closed a single-flight guard that leaked sandboxes under concurrency. **294 merged PRs** into a repo at 873 stars and 329 forks; backends span Anthropic, OpenAI and Ollama.

- Ocopus** — creator · multi-tenant SaaS business platform, NestJS + SvelteKit rohitk06.in/projects/ocopus
- Runs an entire business, not just a till: POS and kitchen display, inventory with auto-reorder, reservations and waitlist, delivery zones and driver tracking, loyalty and campaigns, staff scheduling, franchise hierarchy, GST and accounting exports. 40+ backend modules over 82 Prisma models, with an offline-first PWA point of sale and AES-256-GCM field encryption behind HMAC blind indexes.
 - Exposes that whole surface to agents as **204 MCP tools**, so an assistant reads stock, books a table or places an order through the same contracts and permission model a human operator uses. One permission model, not two.

- discli** — creator · Discord CLI for AI agents, Python + Claude Agent SDK github.com/DevRohit06/discli
- Server, message, thread and event commands behind security profiles, JSONL audit logging and rate limiting, so a confused agent's blast radius is bounded and every action it took can be reconstructed afterwards. Ships on PyPI as discord-cli-agent: **29k+ installs**.

- Aura IDE** — creator · AI cloud IDE with agentic coding, SvelteKit 5 github.com/DevRohit06/aura-ide
- The agent gets its own Daytona sandbox rather than the user's machine, with provider-agnostic model routing through the AI SDK, so swapping models is configuration and not a rewrite.

SKILLS

AI & agents: generative UI and spec-driven rendering, agent runtimes, tool and function calling, MCP tool surfaces, streaming and token budgeting, provider-agnostic model routing, Claude Agent SDK, pydantic-ai, Vercel AI SDK, sandboxed execution with Daytona. I develop with **Claude Code** daily, which is how 294 merged PRs into one platform is a realistic number.

Backend: Python, Node.js, FastAPI, NestJS, Prisma, PostgreSQL, MongoDB, BullMQ, WebSockets, GraphQL, multi-tenant data and permission models.

Frontend: Svelte 5 / SvelteKit, React, Next.js, Astro, Vue, TypeScript, Tailwind, MDX.

Open source: 419 merged PRs across GitHub, 59 public repositories. [astryx-svelte](#), a Svelte 5 port of Meta's Astryx design system whose CI diffs compiled output against Meta's across 1,528 style keys. [Lito](#), a CLI-first docs engine on npm. Hindi i18n contributor to the official [Astro docs](#).

EDUCATION

SMT. CHM College, Ulhasnagar — BSc Computer Science 2023 – 2026