



RTK

Rust Token Killer · open source, Apache 2.0

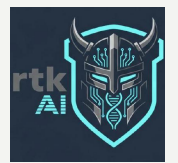
Managing the context window, **from the terminal to a platform**

An open-source Rust binary, on developer machines.

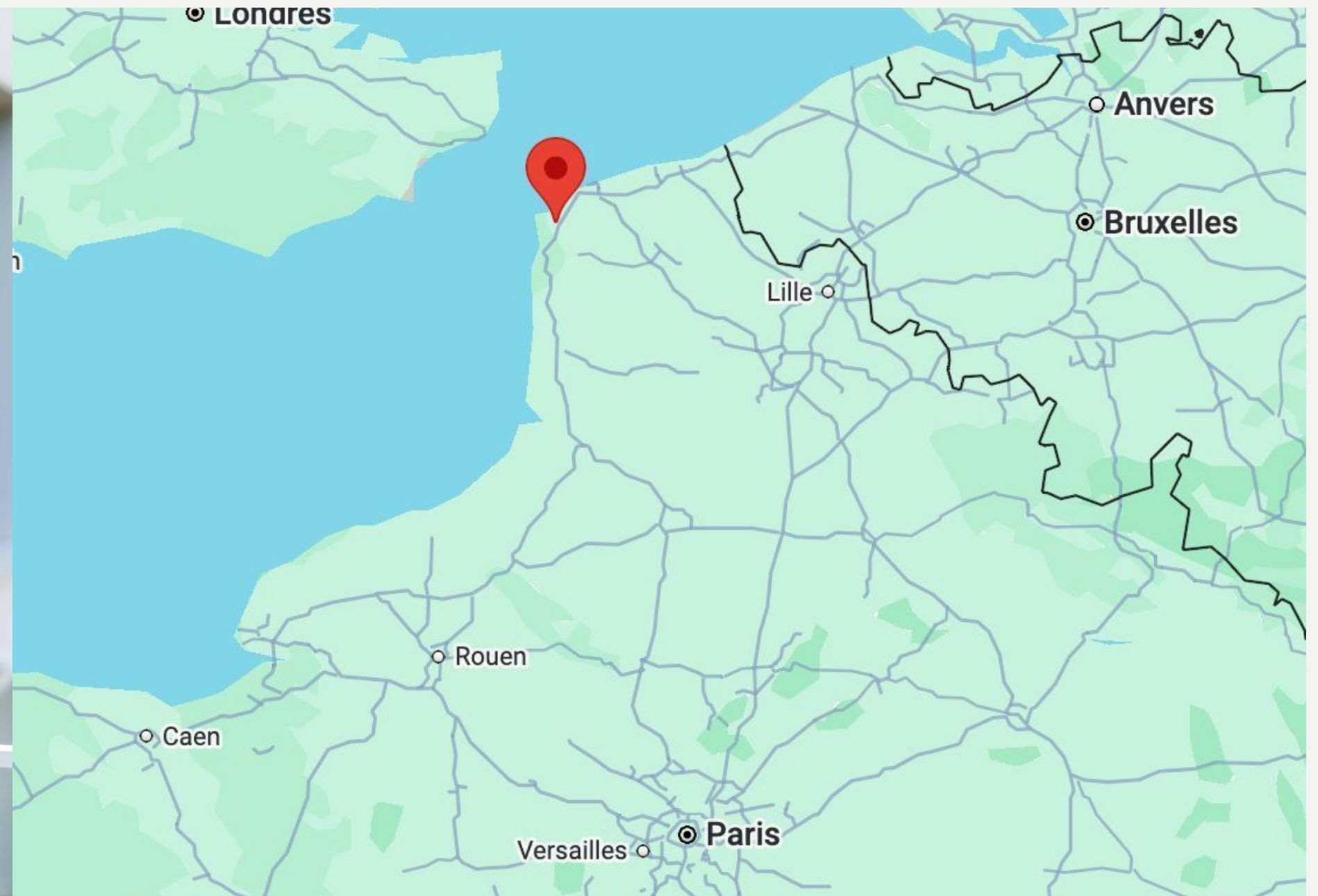
Patrick Szymkowiak

Founder & CEO, RTK AI Labs · Microsoft Certified Trainer

rtk-ai.app



From Wissant to Seattle





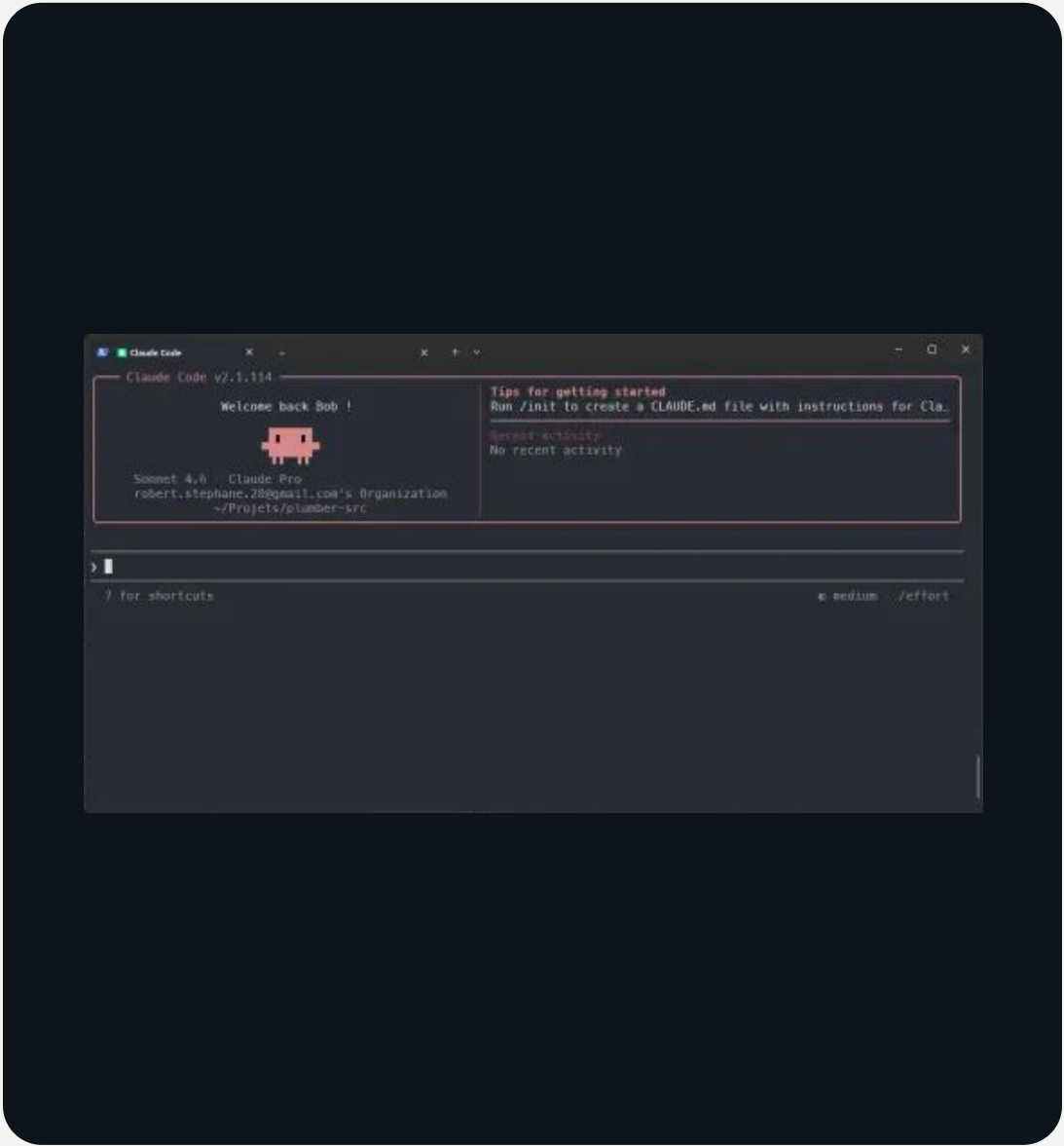
My Tech journey

-
- 01 Engineering degree (Master's level), Telecom & Networks
 - 02 15 years running infrastructure for large companies
 - 03 2 VMworld Europe awards: virtualisation and private cloud
 - 04 Founder, InnovTech (2017)
 - 05 Azure architect, Microsoft Certified Trainer
 - 06 2 years putting AI into production on Azure
 - 07 Founder of 2 companies. Kexa and RTK · 2024–2026
-

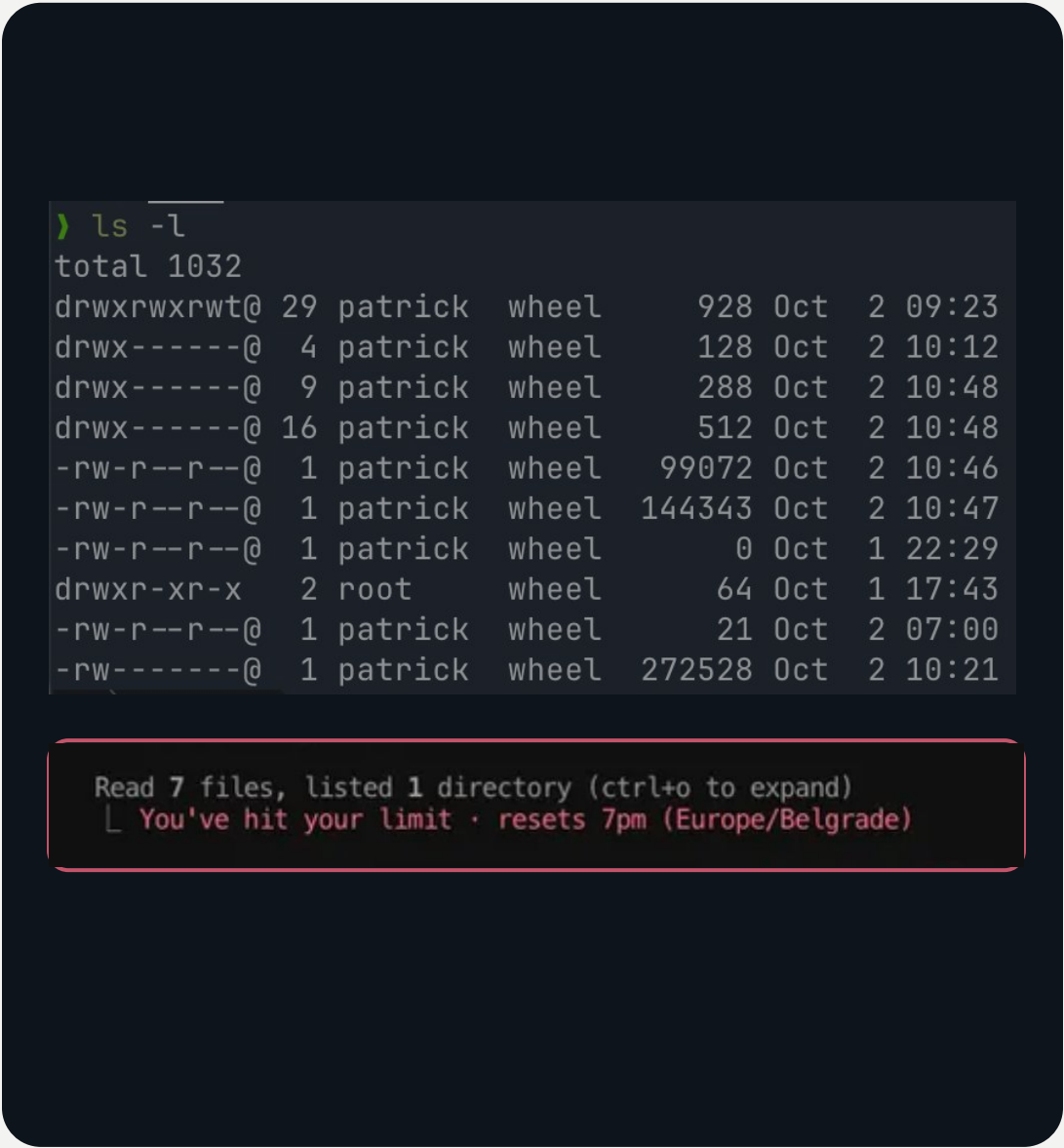
Why RTK?



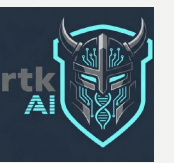
01 July 2025:
I discovered vibe coding with Gemini CLI.



02 September 2025:
I continued with Claude Code.



03 October 2025:
No tokens left on my subscription.
I kept hitting session limits.



RTK beginning...

Without RTK 1,042 lines

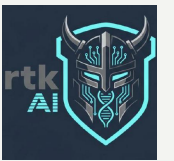
```
$ cargo build
  Compiling version_check v0.9.4
  Compiling syn v2.0.66
  Compiling bytes v1.6.0
  Compiling futures-core v0.3.30
  Compiling proc-macro2 v1.0.85
  Compiling unicode-ident v1.0.12
  Compiling log v0.4.21
  Compiling quote v1.0.36
  Compiling pin-project-lite v0.2.14
  Compiling smallvec v1.13.2
  Compiling proc-macro2 v1.0.85
  Compiling equivalent v1.0.1
  Compiling once_cell v1.19.0
  Compiling proc-macro2 v1.0.85
warning: unused variable: `raw`
--> src/filter.rs:214:9
```

12 lines

With RTK 4 lines

```
$ rtk cargo build
✓ build finished in 32.4s
  · 214 crates compiled
  · 2 warnings: filter.rs:214, tree.rs:87
```

Same command. Same information. One of them fits in the context window.



Same command, same signal, fewer tokens

cargo build went from 1,042 lines to 4. This is command output, not the whole session.

```
$ cargo test
```

~~1,042 lines~~

3 lines

```
$ git status
```

~~38 lines~~

5 lines

```
$ find . -excl. target/
```

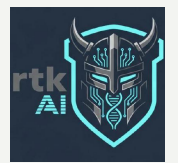
~~312 lines~~

3 lines

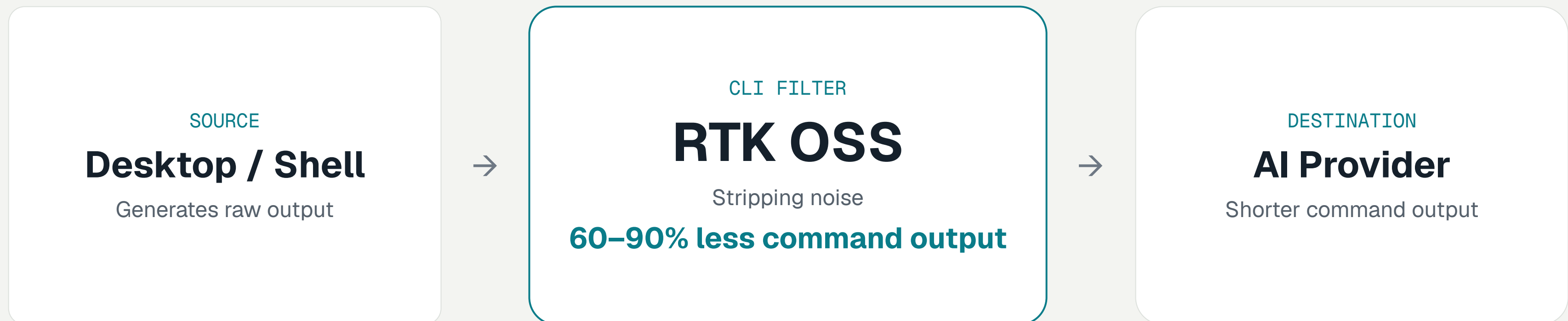
```
$ grep
```

~~217 lines~~

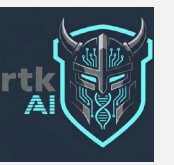
4 lines



Cuts shell command output by 60–90%



Measured on command output, not on the whole session.



A hook between the agent and the shell

```
$ brew install rtk-ai/tap/rtk
$ rtk init --global
✓ PreToolUse hook installed in Claude Code settings.json
✓ every Bash call is now rewritten to its rtk equivalent

$ rtk init --global --agent <name>
  Copilot CLI · Cursor · Codex · Gemini CLI · OpenCode · Windsurf ·
  Cline
```

12

languages

Rust · Go · JS/TS · Python · Ruby · PHP · C# · Java/Kotlin · Scala · Swift ·
Elixir · C/C++

Rust

single static binary

< 20 ms

per intercepted
command

110+

commands covered

17

AI coding tools

One Reddit post, 10 months, 82,017 stars



22 Jan 2026
repo created

first 100 days
0 → 39,045 stars

the trigger
one Reddit post, 360K views

today
82,017 stars · 5,203 forks

Where it runs today

82,017

GitHub stars
5,203 forks

4M+

binary downloads
GitHub release assets, all time

375

releases
shipped since January 2026

163

contributors
Apache 2.0

1,500+

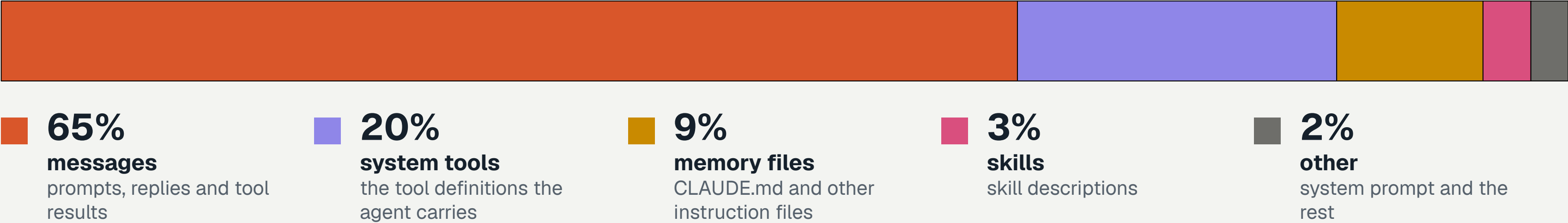
organisations
with two or more stargazers

Starred by developers at

Starred by developers who work at these companies. No commercial relationship.

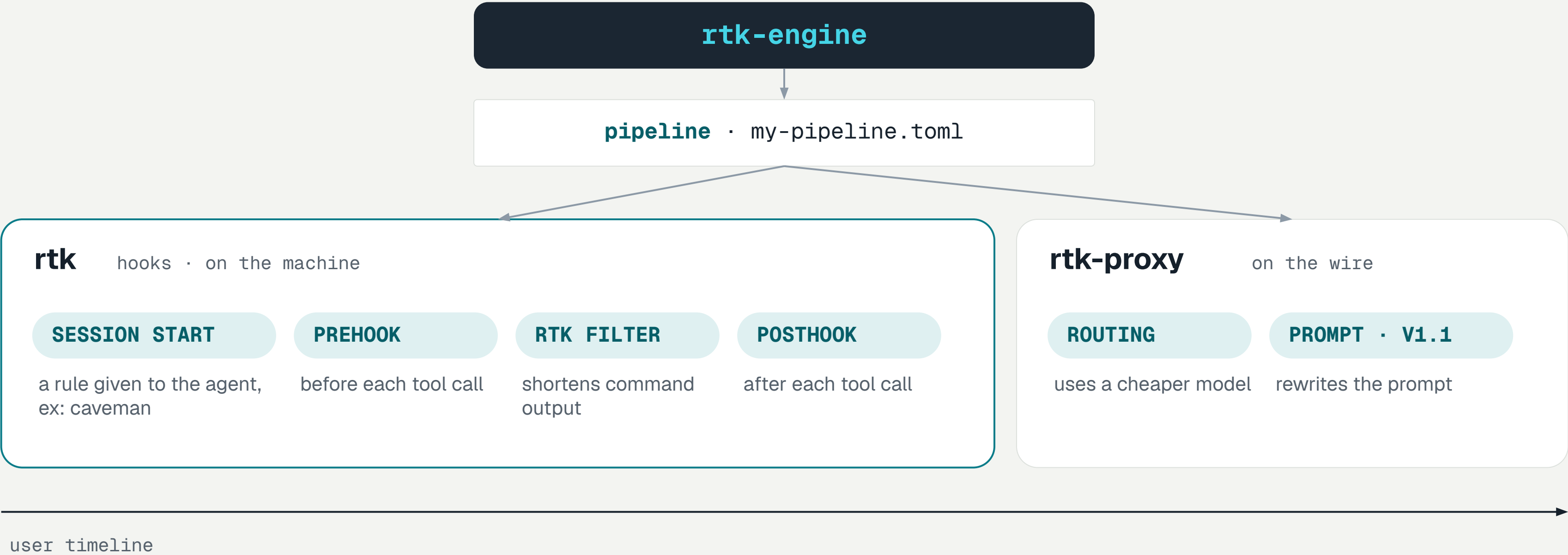
beyond RTK OSS : RTK proxy for all context



RTK Proxy catches 100% of this context

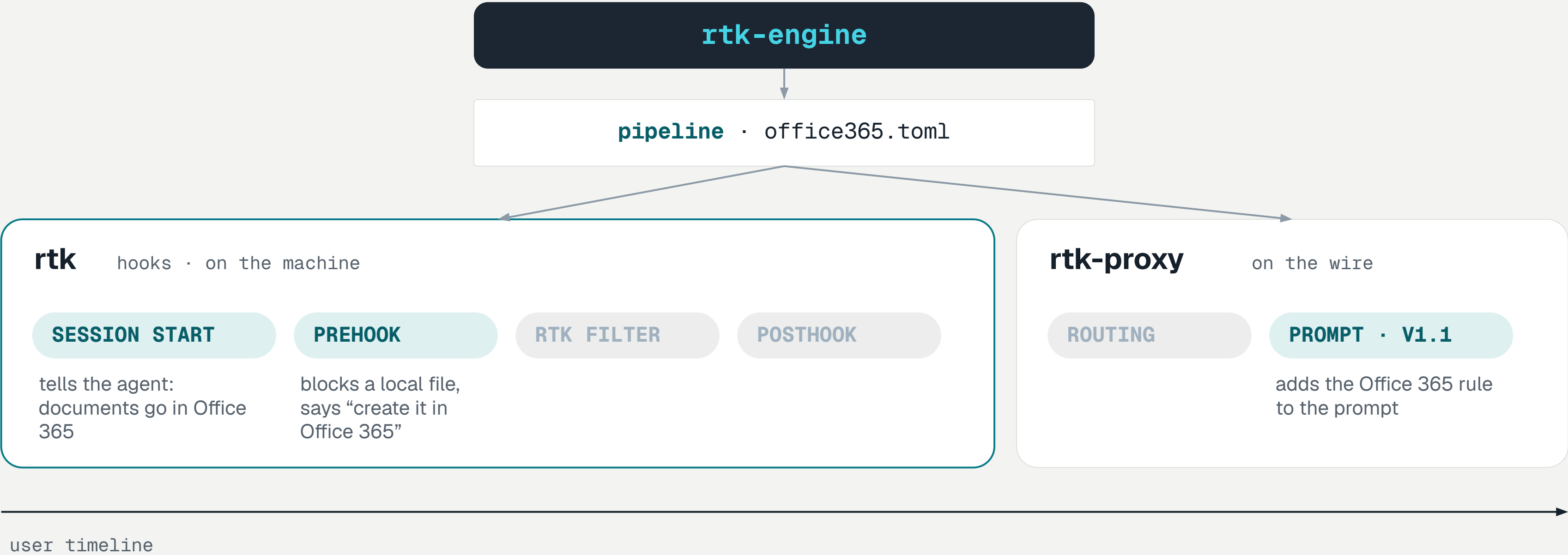
BETA · NOT YET RELEASED

beyond RTK OSS : rtk engine : there's a plugin for that



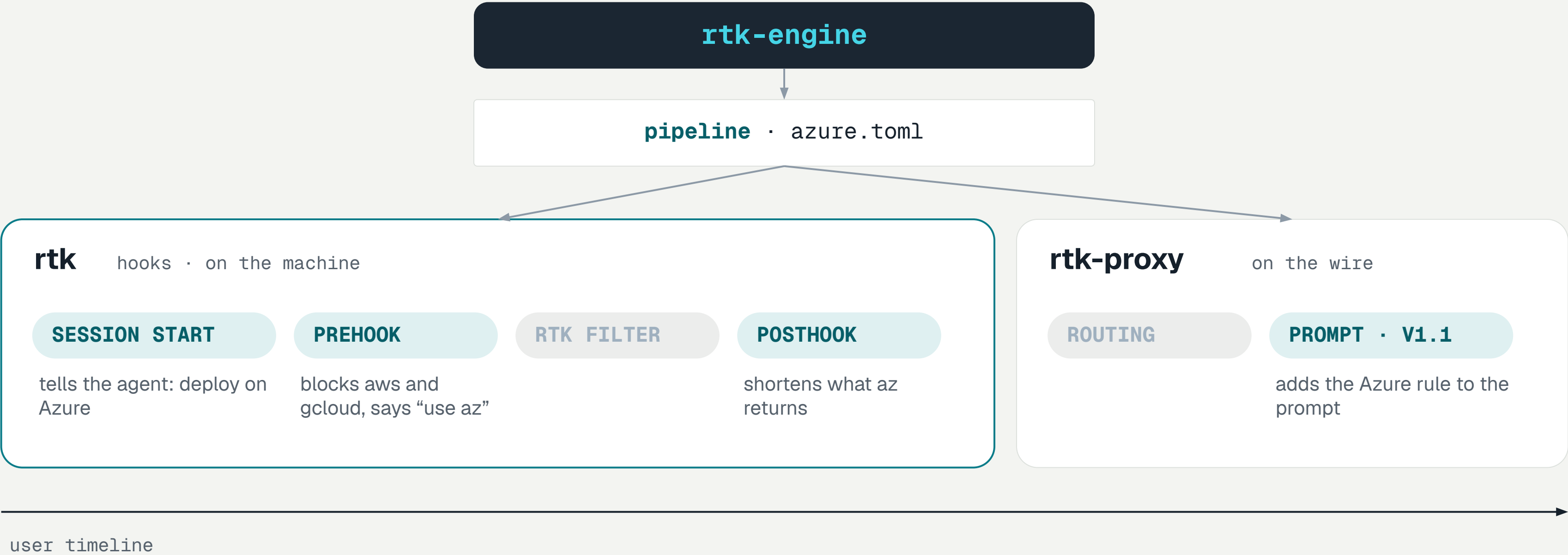
BETA · NOT YET RELEASED

rtk engine : as an Office 365 plugin



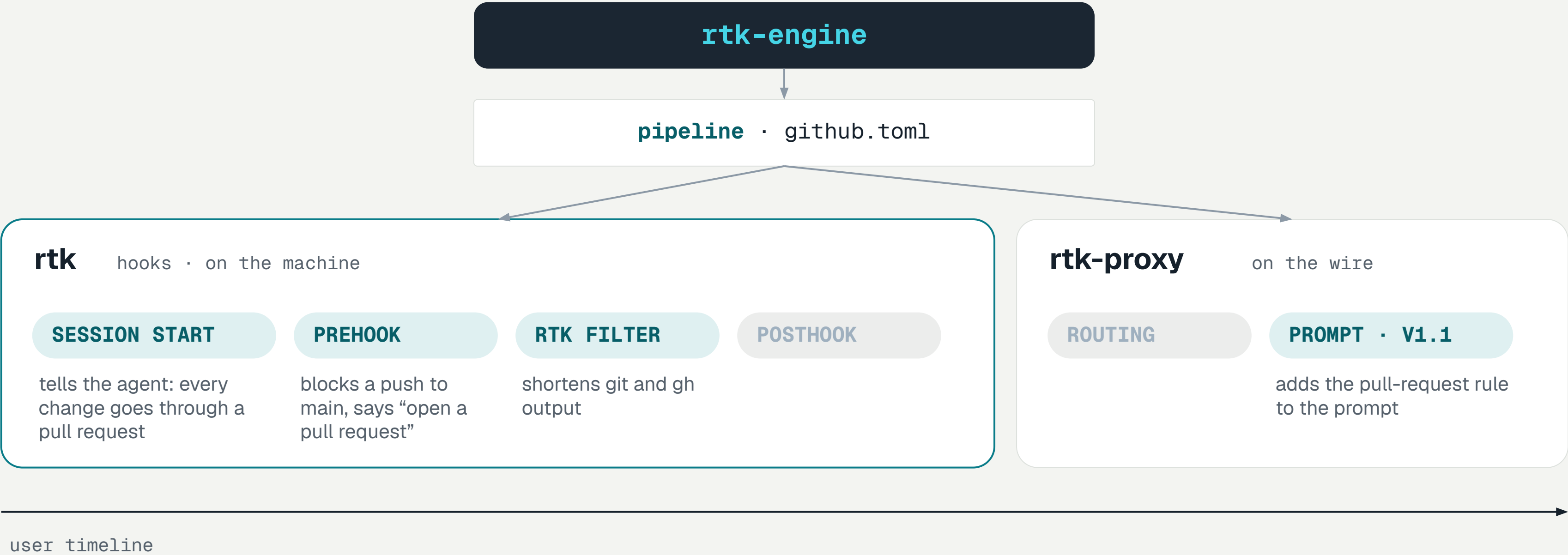
BETA · NOT YET RELEASED

rtk engine : as an Azure plugin



BETA · NOT YET RELEASED

rtk engine : as a GitHub plugin



BETA · NOT YET RELEASED

rtk engine : a skill asks, a rule enforces

A SKILL

Probabilistic



The model decides. Followed most of the time.

VS

AN RTK RULE

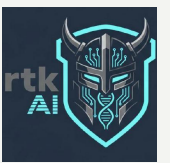
Deterministic



The call is stopped. Applied every time.

RTK Pro

The control plane for large organization



Problems in the real world

UBER

A year's budget, gone in four months

5,000 engineers on Claude Code from December. The 2026 AI coding budget was gone by April; one two-hour session cost \$1,200.

Now capped at \$1,500 per engineer, per agent.

AMAZON

\$1.8M on one project, 860% over

A Claude-based tool to match author details to listings. It never shipped, and nobody noticed the overspend for five months.

Now building automated guardrails on project spend.

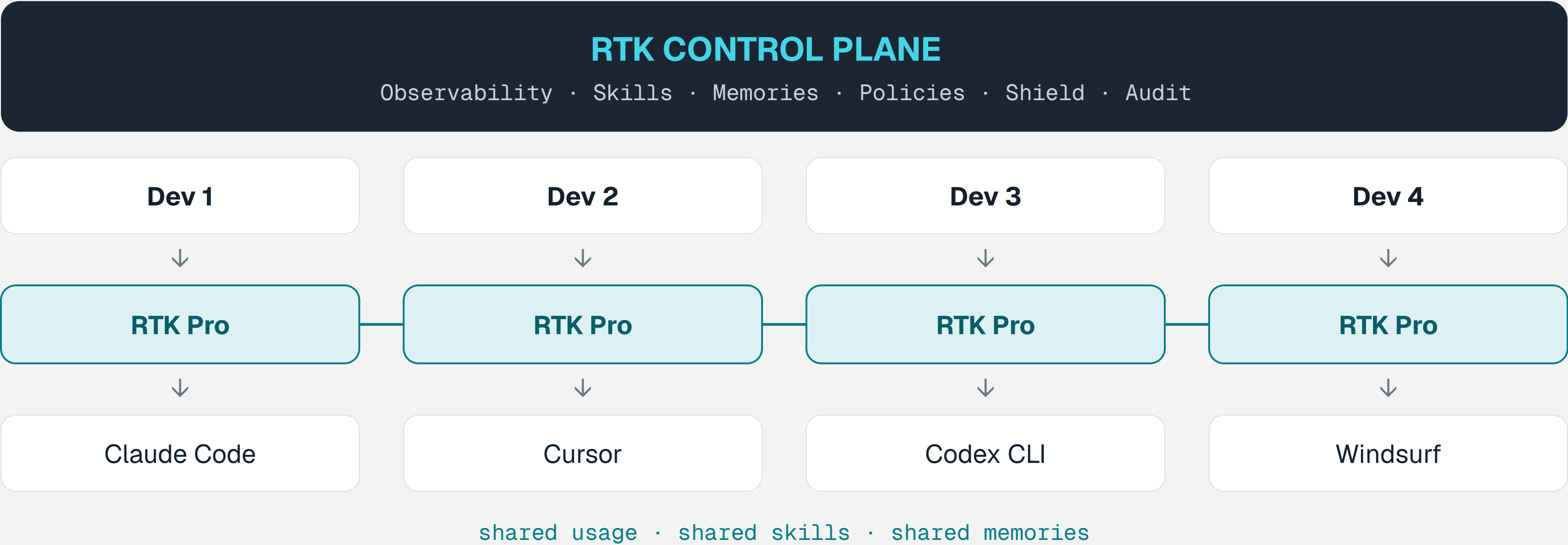
ATLASSIAN

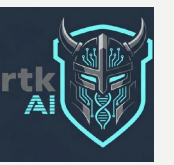
Capped wallets, \$500–2,000

Per-engineer AI budgets set by role, more on request. A cap, not attribution.

A different, unnamed company tripled past \$15M a month.

RTK Pro : one control plane





RTK Pro : a control plane

Secrets

Detected and rewritten in real time with gitleaks patterns: API tokens, keys, private keys. Full coverage on Claude Code, MCP results only on Cursor.

Policies

Allow, ask or block each command, from organisation down to project, before it runs. The local engine is the floor if the cloud is down.

Governance log

Append-only record of every policy and skill decision: allowed, asked or refused, and under which rule. Kept 30 to 90 days.

Skills

Written once, shared across the organisation, enforced rather than suggested. On every plan.

MCP

Servers allowed or denied centrally, from the Teams plan up. Per-call rules come with the v1.0 pipeline.

Memory (roadmap)

ICM: permanent memory for agents, open source and MCP-native. Team sync with provenance is in review.

The team behind RTK

A small team building tools for AI coding agents.



Patrick Szymkowiak
CEO · Founder
github.com/pszymkowiak



Thomas Grauwin
COO · Co-founder
linkedin.com/in/thomasgrauwin



Adrien Eppling
Software Engineer · Co-founder
github.com/aeppling

OPEN SOURCE MAINTAINERS

Nicolas Le Cam
@KuSh

Takayuki Maeda
@TaK08Ki

Florian Bruniaux
@FlorianBruniaux

163
contributors have shipped
code, docs or fixes



Thank you



Your feedback

Patrick Szymkowiak

Founder & CEO, RTK AI Labs

contact@rtk-ai.app rtk-ai.app github.com/rtk-ai/rtk pro.rtk-ai.app