DAILY BRIEF · WHAT AI ENGINEERS ARE BUILDING

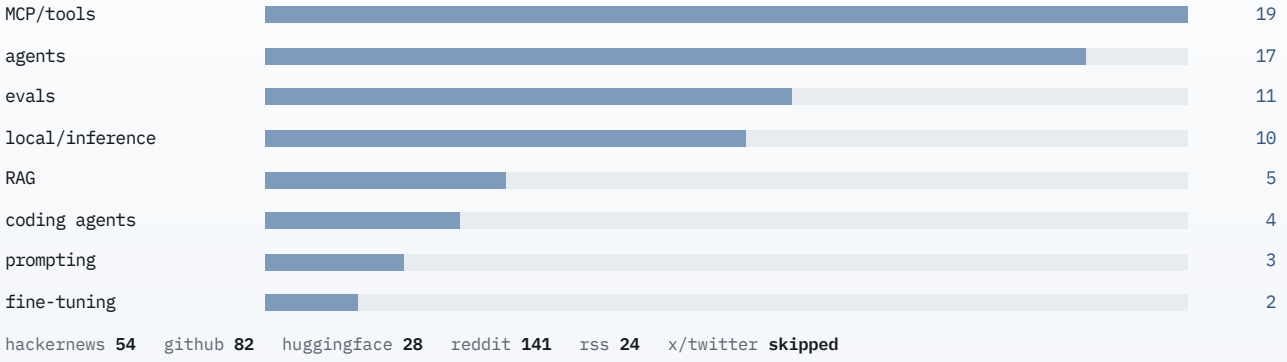
Today in AI engineering 2026-11-04

Daily AI-engineering signal (mechanical fallback — LLM error: 404 Client Error: Not Found for url: https://api.groq.com/openai/v1/chat/complet)

5 deep dives **3** builder stories **15** quick hits **329** items scanned mode: **mechanical fallback**

mcp/tools agents evals local/inference rag coding agents prompting fine-tuning

EXHIBIT 1 · TODAY'S SIGNAL — TOPIC FREQUENCY ACROSS SCANNED ITEMS N = 329



hackernews **54** github **82** huggingface **28** reddit **141** rss **24** x/twitter **skipped**

DEEP DIVES

Show HN: Kairos – Open-source risk and fraud decisioning platform

HN

Hey HN, Ayan here from Finic. We’re launching Kairos, an open source alternative to Sift, Sardine, and other real-time fraud decisioning platforms. Kairos addresses two of the biggest problems facing traditional decisioning tools: a lack of labeled data to train on and dealing with fraudsters that a

Why it matters. Flagged by the ranking heuristics as a real, artifact-backed build.

> **TRY THIS**
> Open the link and skim the repo / paper.

https://github.com/finic-ai/kairos

glincker/awesome-selfhosted-ai

GitHub

Curated list of self-hosted open-source AI: local LLMs, chat UIs, agents, MCP servers, RAG, speech and image generation. Health-checked weekly, dead projects removed automatically.

Why it matters. Flagged by the ranking heuristics as a real, artifact-backed build.

> **TRY THIS**
> Open the link and skim the repo / paper.

https://github.com/glincker/awesome-selfhosted-ai

Show HN: DevBox – Open-source engineering harness from RudderStack

Show HN

Hi, this is Soumyadeb, founder of RudderStack. Today, we are open-sourcing DevBox, our engineering harness we have been using @ RudderStack since April. About 35% of our merged PRs, over 1900 so far, have come from DevBox. Our team still uses Claude and Codex for bigger projects but most small to mi

Why it matters. Flagged by the ranking heuristics as a real, artifact-backed build.

> **TRY THIS**
> Open the link and skim the repo / paper.

https://github.com/rudderlabs/devbox

nvidia/NVIDIA-Nemotron-3.5-Transcribe-30B-A4B

HF model

Trending model on Hugging Face — automatic-speech-recognition, nemo; 1293 likes, 44,939 downloads.

Why it matters. Flagged by the ranking heuristics as a real, artifact-backed build.

> **TRY THIS**
> Open the link and skim the repo / paper.

https://huggingface.co/nvidia/NVIDIA-Nemotron-3.5-Transcribe-30B-A4B

ruvnet/ruvllm

GitHub

Governed neural weight editing and coding model adaptation. Native Rust patches, exact rollback, LoRA, Claude Code routing, and reproducible benchmarks.

Why it matters. Flagged by the ranking heuristics as a real, artifact-backed build.

> **TRY THIS**
> Open the link and skim the repo / paper.

https://github.com/ruvnet/ruvllm

WHAT PEOPLE ACTUALLY BUILT

Show HN: Run Conway's Game of Life in CSS

Show HN

Repository: https://github.com/el-ev/wss This project is heavily inspired by Lyra Rebane's x86CSS (https://lyra.horse/x86css/). So I built a WebAssembly-to-CSS transpiler and spent some time optimizing its output. This is one of the programs it produced: the program runs Conway

> **STEAL THIS**
> Flagged by keyword heuristics as a first-person build write-up — read the post for the details.

https://bucket.svc.moe/uploads/wss/life.html

Show HN: I got tired of GitHub trending, so I built my own

Show HN

I like to see what new projects people are working on in open-source, but I found it hard to find projects or news in any systematic way. I built Shadowgraph so that I could browse the latest in open source development in a more natural way. It ingests repos from github and scrap

> **STEAL THIS**
> Flagged by keyword heuristics as a first-person build write-up — read the post for the details.

https://shadowgraph.io

Show HN: I made a portals physics engine to prove my point

Show HN

I assume everyone has seen the supposedly portal "paradox" where a blue portal comes smashing down on a cube and people argue if it shoots out the orange portal or not. Well, I built a simulation with a set of scenarios to look at, each introducing a new concept, slowly building

> **STEAL THIS**
> Flagged by keyword heuristics as a first-person build write-up — read the post for the details.

https://portalphysics.xyz/up-down

QUICK HITS

exzade/EXZA-Remote GITHUB

Secure AI-powered remote computer access for ChatGPT, Claude and compatible MCP clients.
https://github.com/exzade/EXZA-Remote

abenzerps/Qwen-Image-2.1-Turbo-Quantized HF MODEL

Trending model on Hugging Face — text-to-image, gguf; 5640 likes, 305,860 downloads.
https://huggingface.co/abenzerps/Qwen-Image-2.1-Turbo-Quantized

thefahien/elpis-on-7900xtx GITHUB

Qwen3.8-27B EXL3 + DFlash2 speculative decoding on an AMD RX 7900 XTX (exllamav3-rocm): recipe, patches, benchmarks
https://github.com/thefahien/elpis-on-7900xtx

LiquidAI/LFM3-1.7B HF MODEL

Trending model on Hugging Face — text-generation, transformers; 1453 likes, 54,959 downloads.
https://huggingface.co/LiquidAI/LFM3-1.7B

beipo-ai/NorthCUA GITHUB

Turn intent into work. Open-source Desktop/Web agent harness with long-horizon memory, live evaluation and recoverable batches.
https://github.com/beipo-ai/NorthCUA

deepseek-ai/DeepSeek-V4.1-Flash-1017 HF MODEL

Trending model on Hugging Face — image-text-to-text, transformers; 5225 likes, 221,621 downloads.
https://huggingface.co/deepseek-ai/DeepSeek-V4.1-Flash-1017

wendersoon/radar-cnpj-mcp GITHUB

Servidor MCP e CLI para localizar empresas brasileiras por CNAE e região, utilizando dados abertos de CNPJ da Receita Federal.
https://github.com/wendersoon/radar-cnpj-mcp

capri-com/Capri-AI GITHUB

Capri — Edgerunners-inspired AI agent for your desktop. Works with Claude Code, Codex, Antigravity & more.
https://github.com/capri-com/Capri-AI

mistralai/Mistral-Large-4-1T-A52B HF MODEL

Trending model on Hugging Face — image-text-to-text, transformers; 1571 likes, 8,200 downloads.
https://huggingface.co/mistralai/Mistral-Large-4-1T-A52B

speridlabs/idol HF MODEL

Trending model on Hugging Face — image-to-3d, idol; 557 likes, 14,613 downloads.
https://huggingface.co/speridlabs/idol

sepehr071/iran-jobs-mcp GITHUB

Unofficial read-only MCP server for Iranian job sites (استخدام) IranTalent, Jobinja and Jobvision: job search, companies, reviews, interview
https://github.com/sepehr071/iran-jobs-mcp

ai-gateways/awesome-ai-gateways GITHUB

Awesome AI gateways: open source and commercial LLM gateways, MCP gateways, agent (A2A) gateways, model routers, specs, and guides.
https://github.com/ai-gateways/awesome-ai-gateways

Pika-Labs/Pika-H3-Flex HF MODEL

Trending model on Hugging Face — image-to-video, diffusers; 412 likes, 749 downloads.
https://huggingface.co/Pika-Labs/Pika-H3-Flex

Lightricks/LTX-2.5 HF MODEL

Trending model on Hugging Face — image-to-video, diffusion-single-file; 10557 likes, 2,493,402 downloads.
https://huggingface.co/Lightricks/LTX-2.5

Revolut moves to self-hosted GLM HN

https://twitter.com/guilleflorvs/status/2117379073785696395

Generated 2026-11-04 · Sources: Hacker News, Reddit, GitHub, Hugging Face, arXiv, engineering RSS. Filtered for real, implementable work; marketing and cash-grabs removed. Built by the ai-daily-brief automation · LUCA CONARROE — AI Automation.