

DAILY BRIEF · WHAT AI ENGINEERS ARE BUILDING

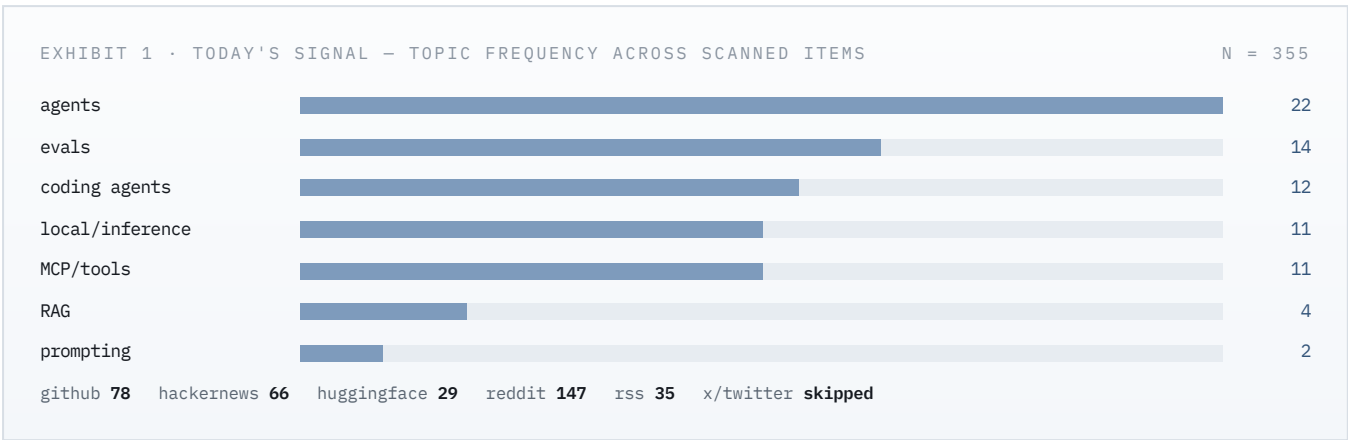
Today in AI engineering

2026-09-24

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 355 items scanned mode: **mechanical fallback**

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



DEEP DIVES

incoai/splash

GitHub

A local inference engine for Apple silicon, built around the model.

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/incoai/splash>

Show HN: How long do I need to work at my salary before I can coast, or retire?

Show HN

I made a financial planning tool in a couple hours that can model various stages of life and potentially guide job searching and work timelines.

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/karmanyaahm/budget-tools/tree/main/fire>

convaiinnovations/laya

HF model

Trending model on Hugging Face — text-classification, transformers; 3244 likes, 0 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/convaiinnovations/laya>

pallavi-shekhar/ai-engineering-interview-questions-company-wise

GitHub

Your Cheat Sheet For AI Engineering Interviews at Top AI Companies - Questions and Answers.

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/pallavi-shekhar/ai-engineering-interview-questions-company-wise>

Show HN: Droid ASC – An On-Demand Android Decompiler, 41–269x Faster Than JADX

Show HN

A few days ago, I was using JADX MCP to analyze my phone, but JADX MCP was just too laggy. I had a large number of APKs to analyze, and nearly every one of them required spinning up a new JADX process. With just two or three of them, my computer ran out of memory. The efficiency was terribly low, so

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/MG1937/ASC>

WHAT PEOPLE ACTUALLY BUILT

Show HN: I built a post-mortem debugger for native Windows x64/x86 crashes

Show HN

Hello HN! I've spent years debugging Windows crashes with tools that were either friendly but limited (e.g. Visual Studio) or powerful but archaic (e.g. WinDbg). I developed patterns and methods for understanding what was going on, and decided to build it into a much more effecti

STEAL THIS

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

<https://www.forensicdbg.com>

Show HN: I made my own scripting language for my game engine

Show HN

(no summary)

STEAL THIS

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

<https://github.com/ArcadeMakerSources/ArcadeMaker/tree/master>

I built a free, open-source desktop workspace for multiple coding agents

r/LLMDevs

I was juggling Codex, Claude Code, OpenCode, and other agent apps. My PC felt slow and I kept losing track of which project or session was where. That became CodeInk: a desktop workspace for Codex, Claude Code, OpenCode, Pi, and ACP agents. It uses your existing agent installs an

STEAL THIS

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

https://www.reddit.com/r/LLMDevs/comments/1wof8z6/i_built_a_free_opensource_desktop_workspace_for/

QUICK HITS

anandi1989/awesome-jev-usecases GITHUB

Evidence-backed index of real-world Jev (TypeSafe AI System One) use cases: repos, patterns, benchmarks, and measured results

<https://github.com/anandi1989/awesome-jev-usecases>

Qwen/Qwen-Image-2.1 HF MODEL

Trending model on Hugging Face — text-to-image, diffusers; 2104 likes, 37,618 downloads.

<https://huggingface.co/Qwen/Qwen-Image-2.1>

prism-ml/Ternary-Bonsai-2-27B-gguf HF MODEL

Trending model on Hugging Face — text-generation, llama.cpp; 1982 likes, 2,991,233 downloads.

<https://huggingface.co/prism-ml/Ternary-Bonsai-2-27B-gguf>

sno-ai/sno-station GITHUB

Sno Station — your Claude Code and Codex working as one squad on your own machine. Shared encrypted memory, agent-to-agent messaging (Reach)

<https://github.com/sno-ai/sno-station>

Jev Decision Index HF SPACE

Trending Space (static, 184 likes). Benchmarks and news on various repros of TypeSafe's Jev

<https://huggingface.co/spaces/multimodalart/jev-decision-index>

XingChen-AGI/Xing4.0-29B-A4B HF MODEL

Trending model on Hugging Face — text-generation, transformers; 1640 likes, 41,923 downloads.

<https://huggingface.co/XingChen-AGI/Xing4.0-29B-A4B>

abenzers/Qwen-Image-2.1-Uncensored-GGUF HF MODEL

Trending model on Hugging Face — text-to-image, gguf; 1526 likes, 575,697 downloads.

<https://huggingface.co/abenzers/Qwen-Image-2.1-Uncensored-GGUF>

Heman10x-NGU/openJev-verdict-2.0 GITHUB

Calibrated 151M Non-Autoregressive Decision Engine beating TypeSafe Jev & Laya on LocalLLaMA/typed-decisions (77.10% acc, 0.0636 Brier, 0.01

<https://github.com/Heman10x-NGU/openJev-verdict-2.0>

zeeshan8281/slo-router GITHUB

SLO-aware LLM inference router with Jev decisions, live queue metrics, counterfactual evaluation, and reproducible latency/cost benchmarks

<https://github.com/zeeshan8281/slo-router>

Show HN: AgentRun: DSL to turn agents into Workflows SHOW HN

Hi HN, I just open sourced the DSL that our harness in grep.ai uses to turn repeatable parts of agent work into workflows. You can combine t

<https://github.com/Parcha-ai/agentrun>

INSANE0777/Awards-mcp GITHUB

Free, open-source MCP server that gives AI coding agents (Claude Code, Codex, Cursor, OpenCode, ZCode...) design inspiration from award-winnin

<https://github.com/INSANE0777/Awards-mcp>

Show HN: BuildABrowser (Java) – Passes Acid2, Has CSS Grid and an Android port SHOW HN

BuildABrowser Browser is an open-source browser and its rendering engine. It runs on Linux, Windows, and (experimental) Android Desktop. The <https://buildabrowser.net/>

Simreal-AI/Simreal-MLBench GITHUB

Innovative externally-scored agentic ML benchmark for next evolution of AI — real competition ground truth, two-submission feedback, attacks <https://github.com/Simreal-AI/Simreal-MLBench>

Comfy-Org/Qwen-Image-2.1 HF MODEL

Trending model on Hugging Face — model, diffusion-single-file; 649 likes, 2,858,923 downloads. <https://huggingface.co/Comfy-Org/Qwen-Image-2.1>

valentynkit/awesome-jev-typesafe GITHUB

Typed decisions with TypeSafe's Jev, the first System One model <https://github.com/valentynkit/awesome-jev-typesafe>
