

DAILY BRIEF · WHAT AI ENGINEERS ARE BUILDING

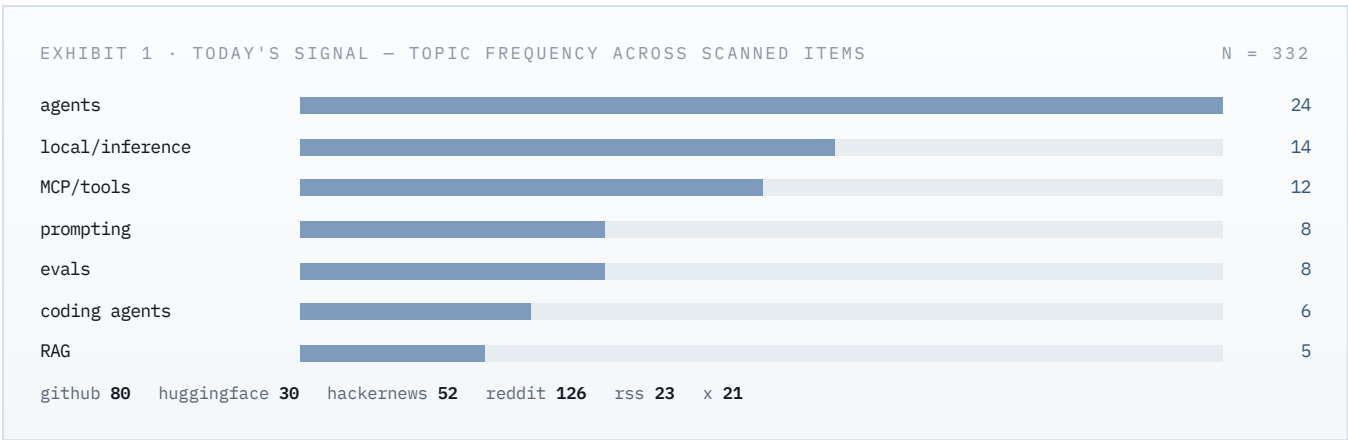
Today in AI engineering

2026-07-19

Developers are building custom AI agents and tools for various applications, including coding, trading, and social media management.

3 deep dives 2 builder stories 10 quick hits 332 items scanned

- ai agents
- llm tooling
- rag security
- model evaluation
- on-device inference



DEEP DIVES

PengZhang64/circuit-framework

GitHub

Peng Zhang built a multi-agent LLM trading research system called Circuit Framework. The system allows users to create and manage trading strategies using LLMs. The project is open-source and available on GitHub.

Why it matters. Circuit Framework demonstrates the potential of LLMs in trading and finance. It provides a platform for researchers and developers to experiment with new trading strategies and models.

TRY THIS

Explore the Circuit Framework repository and experiment with creating your own trading strategies using LLMs.

<https://github.com/PengZhang64/circuit-framework>

- llms
- trading
- finance

vshulcz/deja-vu

[GitHub](#)

Vshulcz built a memory layer for coding agents called Deja Vu. The tool provides features such as search, MCP recall, auto-context, and secret redaction. It is designed to work with various coding agents, including Claude Code and Codex.

Why it matters. Deja Vu addresses the need for efficient memory management in coding agents. It provides a solution for developers to improve the performance and productivity of their coding agents.

TRY THIS

Install Deja Vu and integrate it with your coding agent to improve its memory management capabilities.

<https://github.com/vshulcz/deja-vu>

[coding agents](#)[memory management](#)[productivity](#)

thinkingmachines/Inkling

[Hugging Face](#)

The Thinking Machines team released a trending model on Hugging Face called Inkling. The model is designed for image-text-to-text tasks and has gained significant attention from the community.

Why it matters. Inkling demonstrates the advancements in image-text-to-text models and their potential applications. It provides a valuable resource for developers and researchers working on similar tasks.

TRY THIS

Experiment with the Inkling model on Hugging Face and explore its capabilities for image-text-to-text tasks.

<https://huggingface.co/thinkingmachines/Inkling>

[image-text-to-text](#)[hugging face](#)[trending models](#)

WHAT PEOPLE ACTUALLY BUILT

Show HN: iMessage is the best interface, so I built TypeScript SDK for it

[Hacker News](#)

Jmisilo built a TypeScript SDK for iMessage to integrate it with their personal agent.

How it works. The SDK provides a simple and efficient way to interact with iMessage, allowing developers to build custom interfaces and integrations.

STEAL THIS

When building custom interfaces, consider using existing platforms like iMessage to simplify the development process and improve user experience.

<https://github.com/jmisilo/imessage-sdk>

Show HN: Loris Tunnel – an open-source desktop manager for SSH tunnels

[Hacker News](#)

RangerWolf built an open-source desktop manager for SSH tunnels called Loris Tunnel.

How it works. Loris Tunnel provides a user-friendly interface for managing SSH tunnels, making it easier to connect to remote servers and services.

STEAL THIS

When building custom tools, consider using existing technologies like SSH tunnels to simplify the development process and improve user experience.

<https://github.com/RangerWolf/loris-tunnel-app>

QUICK HITS

KlaatAI/klaatcode GITHUB

Klaatcode is an open-source AI coding agent for the terminal.
<https://github.com/KlaatAI/klaatcode>

PromptPartner/agentsmith GITHUB

Agentsmith is a universal, model-agnostic operating harness for AI agents.
<https://github.com/PromptPartner/agentsmith>

prism-ml/Ternary-Bonsai-27B-gguf HUGGING FACE

Ternary-Bonsai-27B-gguf is a trending model on Hugging Face for text-generation tasks.
<https://huggingface.co/prism-ml/Ternary-Bonsai-27B-gguf>

zeraix/zeraix GITHUB

Zeraix is an open-source local AI workspace for advancing on-device inference.
<https://github.com/zeraix/zeraix>

agentic-commerce-lab/pawl GITHUB

Pawl is a registry-as-ground-truth verification and autonomous research framework for agent-driven development.
<https://github.com/agentic-commerce-lab/pawl>

Show HN: Qwen3.6-35B-A3B on a 16 GB M1 Pro with SSD-streamed MoE HACKER NEWS

This project demonstrates the feasibility of running large models on limited hardware.
<https://github.com/andreaborio/ds4>

thinkingmachines/Inkling HUGGING FACE

Inkling is a trending model on Hugging Face for image-text-to-text tasks.
<https://huggingface.co/thinkingmachines/Inkling>

JesusRS1/stock-trade-finance-api GITHUB

This project provides a stock trade MCP server in Tiingo with stock prices, news, forex, and fundamentals.
<https://github.com/JesusRS1/stock-trade-finance-api>

MemTensor/memmy-agent GITHUB

Memmy-agent is a project that shares a single memory across multiple AI agents.
<https://github.com/MemTensor/memmy-agent>

try1004/FlowLens-AgentOps GITHUB

FlowLens-AgentOps provides observable diagnostics, failure attribution, metrics, and no-key replay for multi-agent runtimes.
<https://github.com/try1004/FlowLens-AgentOps>
