

DAILY BRIEF · WHAT AI ENGINEERS ARE BUILDING

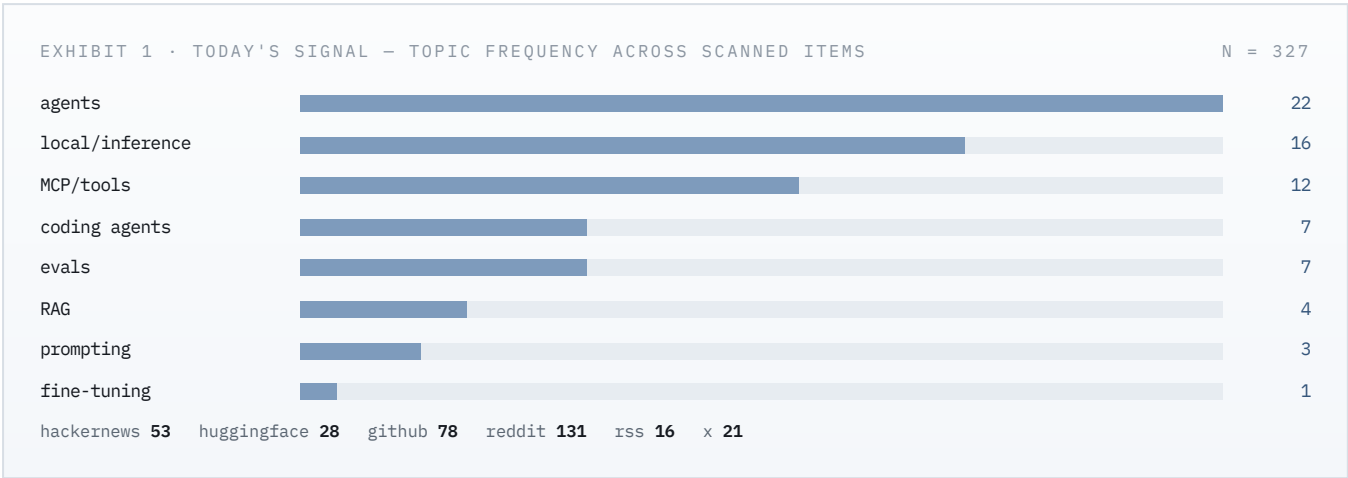
Today in AI engineering

2026-08-02

Advances in AI agent development and LLM tooling dominate today's signal

3 deep dives 2 builder stories 14 quick hits 327 items scanned

llm ai agents transformer models agentic systems



DEEP DIVES

moonshotai/Kimi-K3

huggingface

Moonshot AI has released Kimi-K3, a trending model on Hugging Face that showcases image-text-to-text capabilities with transformers. The model has gained significant attention with 9523 likes and 559,924 downloads. Developers can leverage this model for various applications, including text generation and image understanding.

Why it matters. Kimi-K3 demonstrates the power of transformer models in handling complex tasks like image-text-to-text translation, offering a valuable resource for developers working on similar projects.

TRY THIS

Experiment with fine-tuning Kimi-K3 for a specific task, such as generating text based on image inputs.

<https://huggingface.co/moonshotai/Kimi-K3>

transformer models image-text translation llm

QwenAudio/qwen-audio-agent

[github](#)

QwenAudio has introduced a real-time voice runtime designed for AI agents, enabling them to engage in continuous conversation. This innovation can significantly enhance the interaction capabilities of agents in various applications.

Why it matters. The QwenAudio agent represents a leap forward in real-time voice processing for AI agents, potentially transforming how agents communicate and interact with users.

TRY THIS

Explore integrating the QwenAudio agent into an existing project to enhance voice interaction capabilities.

<https://github.com/QwenAudio/qwen-audio-agent>

[real-time voice processing](#)[ai agents](#)[conversational ai](#)

lordx64/pentestkit

[github](#)

Lordx64 has developed an autonomous multi-agent pentest framework that leverages Kimi K3 for enhanced vulnerability assessment. This framework automates the process of identifying and exploiting vulnerabilities, making it a powerful tool for security testing.

Why it matters. The pentestkit offers a comprehensive approach to penetration testing, utilizing AI agents to streamline the process and improve the efficiency of security assessments.

TRY THIS

Investigate how the pentestkit can be integrated into existing security protocols to enhance vulnerability detection.

<https://github.com/lordx64/pentestkit>

[penetration testing](#)[ai agents](#)[security](#)

WHAT PEOPLE ACTUALLY BUILT

Show HN: I'm a 15 Year Old Wannabe Engineer, This Is a Cycloidal Gearbox I Built

[hackernews](#)

A 15-year-old engineer built a cycloidal gearbox, showcasing their mechanical engineering skills and creativity.

How it works. The gearbox is designed to demonstrate the principles of mechanical advantage and gear ratios, providing a practical example of engineering principles in action.

STEAL THIS

Young engineers can apply similar principles to design and build innovative mechanical systems.

https://github.com/tom-ilan/cycloidal_gearbox

Show HN: Crew, a local collaborative (people and agents) IDE

[hackernews](#)

JamelHammoud and their brother developed Crew, a collaborative IDE that allows people and AI agents to work together on projects.

How it works. Crew enables real-time collaboration between humans and AI agents, facilitating the development of complex projects through shared workspaces and version control.

STEAL THIS

Developers can apply the concept of collaborative IDEs to enhance teamwork and productivity in their own projects.

<https://github.com/JamelHammoud/crew>

QUICK HITS

deepseek-ai/DeepSeek-V4-Flash-0731 HUGGINGFACE

DeepSeek-V4-Flash-0731 is a trending model on Hugging Face for text generation tasks.
<https://huggingface.co/deepseek-ai/DeepSeek-V4-Flash-0731>

OneInterface/stormy-cookbook GITHUB

The Stormy Cookbook provides a comprehensive guide to using the Stormy Social Data API and MCP server for social media data retrieval.
<https://github.com/OneInterface/stormy-cookbook>

bybit-exchange/kaas GITHUB

Kaas is an LLM knowledge-base compiler that turns notes and documents into a queryable Markdown wiki.
<https://github.com/bybit-exchange/kaas>

AnXMuy/AgenticASR GITHUB

AgenticASR is a project focused on refining speech recognition using an agentic approach.
<https://github.com/AnXMuy/AgenticASR>

HezaoHezao/poirot GITHUB

Poirot is a deep research agent kernel designed for those interested in agent architecture.
<https://github.com/HezaoHezao/poirot>

wladimiravila/esp32s3-distributed-ai GITHUB

This project demonstrates distributed LLM inference across ESP32-S3 boards for edge AI applications.
<https://github.com/wladimiravila/esp32s3-distributed-ai>

Evidence-to-Skill – a gate between untrusted sources and agent skills HACKERNEWS

Evidence-to-Skill acts as a gatekeeper between untrusted data sources and AI agent skills, ensuring the reliability of information used by agents.
<https://github.com/Sanexxxx777/evidence-to-skill>

ryan-phq2005h1/github-mcp-server GITHUB

The GitHub MCP Server enables AI agents to interact with the GitHub API, facilitating the automation of GitHub-related tasks.
<https://github.com/ryan-phq2005h1/github-mcp-server>

enowdev/succubus GITHUB

Succubus is a tool for cross-agent coordination, allowing multiple AI agents to collaborate on tasks within a single repository.
<https://github.com/enowdev/succubus>

Oxwi11iam/Medusa GITHUB

Medusa is an open-source AI agentic red-teaming framework that utilizes LangGraph and a Re-Act Architecture for advanced threat simulation.
<https://github.com/Oxwi11iam/Medusa>

gwaghmar/graph GITHUB

This project provides installable graph engineering for various AI coding platforms, including Claude Code and Codex, with features like dependency-graph execution and live reports.
<https://github.com/gwaghmar/graph>

fyfsxkh/TokBrain GITHUB

TokBrain is a knowledge preservation platform designed to sustain wisdom and knowledge over time.
<https://github.com/fyfsxkh/TokBrain>

Show HN: Wienerdog – memory and self-improving skills for Claude Code/Codex HACKERNEWS

Wienerdog enhances Claude Code and Codex with memory and self-improving skills, allowing for more effective and adaptive AI interactions.
<https://github.com/wienerdog-ai/wienerdog/>

Show HN: Aurora – AI Gateway built in Go HACKERNEWS

Aurora is a lightweight AI gateway written in Go, designed to handle API keys, SDKs, and cost tracking for AI projects.
<https://github.com/aurorallm/aurora>

