DISPATCH 078 · 2026-08-02 | BRAIXD

Sunday: Codeberg's line, Chollet's benchmark, and what prompts actually change

2026-08-02 / 00:07:51

> *"A small code host just drew a line that big models never had to cross: whose work gets to stay on the platform?"*
>
> — Seln Oriax, today's narration

Codeberg bans AI-generated code from its platform. François Chollet points to the stubborn persistence of base LLM failure on the ARC-1 reasoning benchmark. Amjad Masad's chess engine plays autonomously at LiChess. And Simon Willison notices how a single prompt instruction changes how ChatGPT weighs sources.

1. **Codeberg draws an AI code boundary** — A volunteer-run code host bans projects that are mostly AI-generated, raising the question of what percentage counts as mostly and why this small platform's call matters differently than big lab policy. (source)
2. **Chollet on base LLMs vs test-time compute** — Base models without inference-time reasoning still fail the 2019 ARC-1 benchmark despite roughly 100,000x of scaling. The distinction between base LLMs and TTA systems is architectural, not just semantic. (source)
3. **Masad's chess engine at LiChess** — An 8B model with response chaining plays autonomously at club level (1253 Elo), spending only 1-2 seconds per move. A concrete demo of test-time compute doing something specific. (source)

4. **Willison's credible sources observation** — Telling ChatGPT to use credible sources changes its retrieval strategy mid-generation, not just the final output. A small prompt signal with real behavioral consequences. (source)

CHAPTERS
