

A Fields Medalist, a PhD chapter, and the week the bar moved

2026-05-09 / 00:31:21

“The lower bound for contributing to mathematics will now be to prove something that LLMs can't prove, rather than simply to prove something that nobody has proved up to now.”

— Lenar Kess, today's narration

A Saturday show that leans into the long reads. Tim Gowers — yes, the Fields Medalist — sat down with ChatGPT 5.5 Pro and an open paper from Mel Nathanson and walked away with a result the original author called "original and clever." We follow that thread, then turn to Mozilla's deeper write-up on the Firefox 271-bug release, Jeff Kaufman on what AI is doing to disclosure embargoes, Anthropic on why constitution training beats demonstration training, and a beautiful pentest story about a critical RCE in React itself. Plus a quieter set of items: Codex in real Chrome, DHH's Copilot review hit-rate jump, a SysMoBench paper on LLM-generated TLA+ specs, AI2's document-routed mixture-of-experts model, and Qwen 35B-A3B running on a 3060.

- [Tim Gowers — A recent experience with ChatGPT 5.5 Pro](#)
- [Mozilla — Behind the Scenes Hardening Firefox](#)
- [Jeff Kaufman — AI is Breaking Two Vulnerability Cultures](#)
- [Anthropic — Teaching Claude why](#)
- [Lachlan — The React2Shell Story \(CVE-2025-55182\)](#)
- [Specula team — Can LLMs model real-world systems in TLA+?](#)

- [OpenAI — Codex in Chrome on macOS and Windows](#)
 - [DHH — Copilot review hit ratio 1/10 to 7/10](#)
 - [r/LocalLLaMA — Qwen 35B-A3B on 12GB VRAM](#)
 - [r/LocalLLaMA — AI2 EMO MoE with document-level routing](#)
 - [METR — Claude Mythos Preview time-horizon evaluation](#)
-

CHAPTERS

- 00:00:04 A Fields Medalist gets a PhD chapter back in two hours
 - 00:04:37 Mozilla unhides 12 of the 271 bugs
 - 00:08:21 AI is breaking two vulnerability cultures
 - 00:11:12 Anthropic on why teaching the constitution beats teaching the answer
 - 00:14:26 React2Shell — the bug an AI pipeline did not find
 - 00:18:33 SysMoBench — when LLMs recite Raft instead of modeling Etcd
 - 00:22:34 Codex moves into your real Chrome
 - 00:25:18 DHH on Copilot crossing some threshold
 - 00:27:03 AI2's EMO and Qwen on a 3060
 - 00:29:17 METR's measurement saturation problem
-