

Where the Goblins Came From, BioMysteryBench, and a Language for Machines

👉 GSV NEVER TALK ABOUT GOBLINS

2026-04-30 · 00:23:39

“Style drift compounds across model generations. The fix doesn't live in your system prompt — it lives in your evals.”

— LENAR KESS, TODAY'S NARRATION

OpenAI publishes a post-mortem on why GPT-5.1 wouldn't stop talking about goblins. Anthropic claims Claude solved 30% of bio problems that stumped expert panels — and an immunologist on X explains what's wrong with that framing. Mistral ships a 128B dense model in a year that has otherwise gone all-in on MoE. IBM's Granite 4.1 8B trades blows with a 32B MoE. Sam Altman gates a frontier cybersecurity model behind a defender ecosystem. WebSockets quietly become the new agent-loop bottleneck-killer. Anthropic's introspection adapters and Qwen's Sparse Autoencoders show up the same week. And a small project called Vera asks the obvious question nobody else is asking: what if you designed a programming language for machines to write?

CHAPTERS

- 00:00:04 Cold open — Never talk about goblins
- 00:03:49 Mistral ships dense, against the grain
- 00:06:41 Granite 4.1 — the 8B that punches up
- 00:08:44 BioMysteryBench — and the immunologist who pushed back
- 00:12:02 Two interpretability artifacts in one week
- 00:14:32 GPT-5.5-Cyber and gated frontier access
- 00:17:59 WebSockets and the new bottleneck

00:20:12 Vera — and the question nobody else is asking

CANONICAL

<https://braid.opentangle.com/episodes/2026-04-30.html>