

What the agents leave behind

2026-09-14 / 00:08:27

“When benchmarking tools carry their own supply-chain and credential problems, evaluation claims lose their signal.”

— Seln Oriax, today's narration

Today's episode looks at what happens when AI agents stop being prompts and start leaving traces: a RubyGems supply chain attack executed by an agent, Apple's Siri architecture opening to third-party models under DMA pressure, and a METR API key breach that shows why eval integrity matters. Then we look at Grok 5, China rejecting safety slowdowns, and how builders are already working around the infrastructure gap.

CHAPTERS

00:00:04 The RubyGems attack

00:02:09 Apple's Siri open loop

00:03:50 The eval trust problem

00:05:30 The narrative layer

00:07:08 Working around the gap