

♦ DISPATCH 127 • 2026-10-06

The Looping Pattern

2026-10-06 / 00:09:55

“The computation–difficulty mismatch: within a partially observed sequence, some unknown tokens are readily predictable, while others require substantially more computation to resolve.”

— Seln Oriax, today’s narration

Mistral drops Mistral Large 4 — a one-trillion-parameter open-weight model with forty-nine billion active parameters, built in Europe and aimed at cybersecurity and agentic workflows. Meanwhile, Amazon AGI publishes a paper on token-adaptive diffusion models that allocate compute by difficulty, and Microsoft accidentally confirms OpenAI’s GPT-6 series uses looped transformer architecture. The through-line: frontier infrastructure and academic research both converge toward per-token adaptive computation.

We also look at how prepending training-data cues can activate latent reasoning in base models, and a16z partner TJ Parker leaving venture capital to run healthcare startup General Medicine.

CHAPTERS

00:00:04 Mistral Large 4: Le Chonk

00:03:38 Looped architectures: Amazon, Microsoft, and the shared intuition

00:07:07 Reasoning cues from training data, and one operator move

