◆ DISPATCH 146 · 2026-09-14 | GSV THE LEASE OUTLIVES THE PLEDGE

We Do Not Believe We Need to Wait

2026-09-14 / 00:27:24

> *"A six-year lease is a forecast with a signature on it."*
>
> — from this episode's transcript

- Lenar Kess
- Damra Vol

The three largest American labs spent the weekend agreeing that capability is moving too fast. Monday was the first trading session after, and it put a price on the agreement — then every government that would have to enforce it declined, and Anthropic signed a $13.7 billion compute lease running six years out.

- The Guardian — SoftBank down 13%, the Kospi off 3%, with the Nasdaq queued to follow. Nothing broke over the weekend except the consensus about pace.
- Sam Altman — welcomes a federal framework but does "not believe we need to wait" for an antitrust exemption or a law. That clause decides whether this is a commitment or an opening position.
- The Verge — Microsoft published a 37-page humanist AI code of conduct, the only actual artifact anyone produced out of the weekend.
- Axios — Trump calls the warnings the work of "negative forces," Speaker Johnson wants a summit, and the Democratic proposals range from a subpoena-powered select committee to twenty-year prison terms.
- The Information, via Techmeme — Anthropic signed a $13.7B, six-year compute lease with Rum Group, which operates Rumble and hosts Truth Social, for a Georgia data center that doesn't exist yet.

- Financial Times, via Techmeme — Anthropic told investors it will be profitable a second straight quarter at 80%+ gross margins, before partner revenue sharing and training costs.
- Reuters, via Techmeme — Samsung and SK Hynix refused to prepay roughly $18.7 billion of Korean chip-cluster power bills, citing uncertainty about long-term demand.
- DropVLA (preprint) — poisoning 0.31% of training episodes forces a chosen robot action to fire on command with a 98%+ success rate and no visible degradation on the actual task.

SEGMENTS

Transcript

1. Lenar Kess 00:00:04

   Picture a trading desk in Seoul on Monday morning. Nothing had broken over the weekend. There was no fab fire, no supply shock, and nothing new out of earnings or the export rules. The only fresh input since Friday's close was a few essays and a handful of posts on X. By the end of the session SoftBank was down thirteen percent and the Kospi off three. SK Hynix, Kioxia, Minimax and Z.ai all dropped more than five. The Guardian's markets desk had the Nasdaq queued to follow. So let's start with a pricing question rather than a philosophy question — what did the market actually reprice?

- theguardian.com
- techmeme.com
- x.com
- techmeme.com
- cnbc.com
- techmeme.com
- theverge.com
- theguardian.com
- reddit.com
- reddit.com
- cohere.com
- axios.com
- aljazeera.com
- x.com
- techmeme.com
- cnbc.com
- nbcnews.com
- techmeme.com
- theguardian.com
- techmeme.com
- x.com
- techmeme.com
- techmeme.com
- clashreport.com
- science.org
- cnbc.com
- techmeme.com
- techmeme.com
- theguardian.com
- techmeme.com
- techmeme.com
- techmeme.com
- arxiv.org
- arxiv.org
- arxiv.org
- x.com

- x.com
- youtube.com
- x.com
- spectrum.ieee.org

2. Damra Vol 00:00:42

   Not capability. Nothing about the models changed between Friday and Monday. The three companies with the most to gain from the next capability jump all said in public that they'd like the next capability jump to arrive later. So what got sold was the <emphasis>schedule</emphasis>.

3. Lenar Kess 00:00:58

   And the names that fell hardest are the ones furthest from the models. SoftBank is a holding company with an OpenAI position. SK Hynix and Kioxia make memory. The Kospi is heavy with suppliers to people who buy accelerators. Nobody sold Anthropic — you can't. They sold the people who sell shovels.

4. Damra Vol 00:01:18

   Which tells you the market read the pledge as a demand-timing statement rather than a safety statement. Though I'd bound how much weight that carries. One Asian session, mostly in illiquid supplier names, on a Monday, with the US market not open yet when those pieces were filed. That's a reaction, not a verdict.

5. Lenar Kess 00:01:38

   Fair. Let's go to the text, because the text is where this gets interesting. Sam Altman posted Sunday night. He writes — quote — "The world deserves confidence that American companies developing increasingly capable AI will act responsibly, especially as the trajectory of progress has steepened. Every frontier lab must deliver on this, and there is no reason any of us should come to work if we cannot." And then the part Techmeme pulled into its headline: he welcomes a federal framework, but doesn't believe they need to wait for an antitrust exemption or a law to pace development responsibly.

6. Damra Vol 00:02:13

   That second clause is the whole document. Over the weekend the open question was whether coordinated pacing between competitors needs legal cover. Can you even sit in a room with your rivals and agree to slow down without a lawyer between you? Altman just answered: no exemption needed, we're doing it anyway. So either he has a theory of why this doesn't count as coordination, or he's decided the enforcement risk is smaller than the reputational risk of waiting.

7. Lenar Kess 00:02:41

He also gave CNBC the sentence that will follow him around — "we could lose control." That's an OpenAI chief executive, on the record, in a business-channel interview.

8. Damra Vol 00:02:52

   And Satya Nadella came in on the same side with different vocabulary. He said Microsoft welcomes the — quote — "deliberate pacing needed to get alignment right," and grounded it this way: "Any pursuit of superintelligence has to be grounded in the core principle that if the AI we build is not helping humanity and under human control, it's not worth pursuing." But he's the only one who attached a document to it.

9. Lenar Kess 00:03:18

   The 37-page humanist AI code of conduct. Tom Warren has it at The Verge, published today, covering Microsoft's own MAI models. The headline line is "people matter more than AI."

10. Damra Vol 00:03:32

    Thirty-seven pages is the only countable thing anybody produced out of this weekend. Everyone else has posts. [chuckle] I'm not saying a PDF is governance. But four companies agreed on a principle, and exactly one of them wrote it down and scoped it to named model families. That's the version you can hold somebody to later.

11. Lenar Kess 00:03:52

    Alright — here's the route today. We stay on the pacing call for one more segment, because the objection to it is more coherent than it looks. Then Washington and Beijing, both of which said no inside forty-eight hours. Then Anthropic's thirteen-point-seven-billion-dollar compute lease, which is a strange thing to sign the same weekend you ask the industry to slow down. After that, a misuse claim involving Claude Code and missile guidance. Then the data center siting fight, which picked up a price tag today. We close on two robotics preprints, Grok 4.8, and a company doing thirty million dollars a year on somebody else's models.

12. Damra Vol 00:04:30

    The objection showed up almost instantly, and it showed up in three rooms at once. On the Anthropic subreddit, somebody posted the blunt version — why are Dario and Altman telling <emphasis>us</emphasis> to slow down AI development? They're the ones developing it. What are we supposed to do about it?

13. Lenar Kess 00:04:47

That's the naive-sounding question that isn't naive at all. The OpenAI subreddit dressed it up and called it the Doomsday Moat. The claim there is that the labs want Washington to stop the clock at precisely the moment they're ahead. The people who lose are open-weight projects and anyone who can't afford a compliance department.

14. Damra Vol 00:05:07

    And Cohere put its name on the same objection in a blog post — "Who gets to define the rules for AI?" — which made Hacker News overnight. Same argument, from an incumbent-adjacent lab rather than a message board, which makes it hard to wave off. Cohere isn't a doomer and isn't an accelerationist. They're a company that would have to live inside whatever rules the three biggest labs negotiate.

15. Lenar Kess 00:05:31

    The sharpest version came from inside the administration's orbit, though. David Sacks — Trump's former AI czar — has a one-line answer, and Axios quotes it plainly: slow yourselves down. If OpenAI and Anthropic truly believe the frontier is moving too fast, they can decelerate unilaterally, today, without Washington forcing anyone else to.

16. Damra Vol 00:05:53

    [tsk] And nobody has answered that in public yet. The reply I'd expect is that unilateral deceleration hands the frontier to whoever doesn't decelerate, and that's worse. The argument holds up — it just only works if somebody eventually binds the other guy. So you're back to needing law.

17. Lenar Kess 00:06:12

    Nick Robins-Early's piece in the Guardian puts the pressure in context, too. He reports that the proposal came after public uproar over Anthropic researchers repeatedly warning that AI could kill all of humanity by 2030. So the sequence matters. The warning was in the world first, the uproar followed, and the pacing proposal came after that.

18. Damra Vol 00:06:33

    Which means you can read it two ways without either being cynical. One: the labs believe the thing their own researchers published and are responding to it. Two: the labs are absorbing a public-opinion problem by getting in front of it. I think both are true at once, and people keep assuming those are exclusive. Companies do the right thing for mixed reasons constantly.

19. Lenar Kess 00:06:56

There's a narrow test here, and it has nothing to do with motive. Does any of these four produce a specific capability threshold with a date attached, or name a third party who gets to check? Yesterday the strongest commitment in the pile was furniture with no recipient. Today it's four companies agreeing plus one code of conduct. That's movement, but it's still not a mechanism.

20. Lenar Kess 00:07:19

    So the labs asked for coordination on Saturday. By Monday, every government that would have to supply it had declined. Trump was in Ireland on Sunday and dismissed the warnings as the work of — quote — "a lot of negative forces ... bringing up things that won't happen," and reduced the whole thing to four words: "Whoever wins AI wins."

21. Damra Vol 00:07:40

    Zachary Basu's Axios piece points out the odd part. The negative forces in question are now the chief executives of Anthropic, OpenAI, SpaceX's AI arm, and Google. The administration has spent two years championing exactly those companies.

22. Lenar Kess 00:07:57

    Speaker Johnson took the Sacks position on Meet the Press. He rejected an emergency moratorium. He said he'd recall the House and hold a vote — quote — "if we had the solution" — and warned Congress against writing rules it doesn't understand yet. His preferred first move is a summit: Trump, congressional leadership, and the lab chiefs in one room.

23. Damra Vol 00:08:18

    A summit is what you schedule when you want the appearance of motion without a bill number. Although — I'll be fair to Johnson — "don't legislate a thing you can't describe" is a defensible position, and it's the same position a lot of engineers would take.

24. Lenar Kess 00:08:32

    The Democratic side is where it gets strange, because there's urgency and no shared doctrine. Jeffries has his caucus meeting tomorrow to work out what they even want. Obama is privately pushing the party to make this a top priority. And the proposals on the table don't converge at all.

25. Damra Vol 00:08:49

    Read the spread out, because the spread <emphasis>is</emphasis> the story. Bernie Sanders and Greg Casar want to ban superintelligence outright, pause advanced development, and put up to twenty years in prison behind it. Ruben Gallego wants a bipartisan select committee with subpoena power. Chris Van Hollen wants OpenAI to open its models to federal safety review. John Thune and

Amy Klobuchar want something much narrower — companies test for catastrophic risk and report what they find.

26. Lenar Kess 00:09:19

Four proposals, and only one of them resembles what the labs actually asked for. Thune–Klobuchar is basically the pacing proposal with a filing requirement bolted on.

27. Damra Vol 00:09:29

And it's the only one that could survive a veto pen, which Axios flags as the reality check on all of it. Their conclusion is the number I keep coming back to — federal consensus is almost unimaginable this year, and probably through early 2029, regardless of who controls Congress. That's a three-year forecast of nothing happening.

28. Lenar Kess 00:09:50

There is one crack, though. Mike Rogers, the Michigan Republican Senate nominee, a Trump ally running in a swing state, said Amodei's warning convinced him — quote — "we must slow down development" so government can catch up. And Josh Hawley is running an investigation into existential risk. That's the coalition that would actually matter if it grew.

29. Damra Vol 00:10:11

The one proposal I've seen with a working mechanism came from outside government entirely. Robin Hanson spent the weekend arguing for mandatory liability insurance — make labs carry coverage against the harms, and let underwriters price the risk. It's the only design where somebody with money on the line has to form an opinion about how dangerous a specific model is, which is more than any of the congressional versions manage.

30. Lenar Kess 00:10:36

Insurance also fails in an obvious way: nobody can price a tail nobody has observed. But I'd rather argue about an actuarial table than about whether a summit should happen.

31. Damra Vol 00:10:47

Beijing's answer took about a day. The Foreign Ministry called the warnings — quote — "fearmongering, confrontation, and vicious competition" that "serve the interests of no one," and Chinese state media added that everyone should be working toward AI openness and inclusivity. NBC's Mithil Aggarwal has the plainer version of what officials there actually think, which is that a coordinated American slowdown is a containment play aimed at them.

32. Lenar Kess 00:11:15

The novel piece isn't the Foreign Ministry, though. Eleanor Olcott at the Financial Times reports the Ministry of State Security issued its first-ever public statement on AI, warning that it poses risks to China's political and social security, including its cyber defenses.

33. Damra Vol 00:11:31

For anyone reading the geopolitics, that's the line to hold onto. The diplomatic arm calls the risk talk a scam. The security service says the danger is genuine and aimed squarely at China. Those are two different Chinese positions published in the same week, and the second one is closer to what the American labs are saying than either government would like to admit.

34. Lenar Kess 00:11:53

And separately, Dan Milmo and Robert Booth report that the UK's cross-party joint committee on human rights — MPs and Lords together — wants a new regulatory framework, an independent oversight body, and legislation. They name public face-scanning and deepfakes as the immediate threats. Different jurisdiction and a different report, in the same week.

35. Lenar Kess 00:12:15

Now the lease. The Information reports — sourced to a person familiar, so hold it loosely — that Anthropic signed a computing deal worth thirteen-point-seven billion dollars over six years with Rum Group, leasing capacity from an upcoming data center in Georgia. Rum operates Rumble and hosts Truth Social.

36. Damra Vol 00:12:34

And Rum became a compute company by buying one. Susan Zhang and a few others were walking through the corporate history overnight. Rumble acquires Northern Data, the German outfit, relaunches as Rum Group, and months later signs a nearly fourteen-billion-dollar lease with an American frontier lab. That's a video platform turning into a neocloud in under a year.

37. Lenar Kess 00:12:57

The political ownership is the part everyone will grab, and I understand why. But the more useful detail is the tense. The Georgia data center is <emphasis>upcoming</emphasis>. It isn't built. Anthropic committed six years of spend to capacity that currently exists as a site plan and a power interconnect application.

38. Damra Vol 00:13:16

A six-year lease is a forecast with a signature on it. And the forecast says: we will want enormously more compute in 2029 and 2030 than we want now, and we're willing to be wrong about that in public, at thirteen-point-seven billion dollars of exposure, rather than risk not having it.

39. Lenar Kess 00:13:34

Which is hard to square with the weekend. The same company whose chief executive asked the industry to pace the frontier just signed a six-year forward commitment to the inputs that make the frontier move.

40. Damra Vol 00:13:46

That's not automatically a gotcha, though, because there's a reading where it holds together. Pacing capability improvement and buying serving capacity aren't the same purchase. If Claude usage keeps compounding, you need that compute whether or not you train a bigger model. [pause] But nobody has said that's what it's for, and the lease term is six years — long for inference planning, and about right for training.

41. Lenar Kess 00:14:10

Put it next to the other Anthropic number from the same twenty-four hours. The Financial Times reports the company told investors it will be profitable for a second straight quarter, with gross margins above eighty percent. The qualifier is the whole sentence though: that's before partner revenue sharing, and before training costs.

42. Damra Vol 00:14:28

Eighty percent before training costs — and then you sign a lease sized for training. Those two disclosures are going to the same audience, investors ahead of an initial public offering, and they point opposite directions. One says the unit economics work. The other says we're about to spend a very large amount below the line that makes them work.

43. Lenar Kess 00:14:48

And the third number of the day fits the same pattern from the other lab. Janice Huang at Bloomberg reports SoftBank closed a two-year loan facility to support its OpenAI investment. About twenty banks committed eleven-point-eight-seven billion dollars against a ten-billion-dollar target. Oversubscribed.

44. Damra Vol 00:15:07

Twenty banks competing to lend against an OpenAI stake, and then on Monday the borrower's equity drops thirteen percent because its counterparty's peers said they'd like to go slower. If you wanted a single picture of how much of this is financed rather than funded, that's it.

45. Lenar Kess 00:15:24

A harder one. There's a report circulating, carried by Clash Report and picked up on Hacker News, attributing to Anthropic the claim that Houthi-affiliated operators used Claude Code to develop missile guidance software.

46. Damra Vol 00:15:38

Let's be precise about the chain before we say anything about the substance, because the chain is thin. A secondary outlet reporting a claim by the vendor, seventeen points and thirteen comments on Hacker News, and no primary threat-intelligence document in front of us. I don't want to treat the phrase "missile guidance software" as confirmed.

47. Lenar Kess 00:15:58

Agreed, and the phrase arrives with no definition attached. Guidance software could mean a Kalman filter someone asked for help debugging. It could mean terminal-phase targeting. Those are enormously different claims, and both would get reported with the same four words.

48. Damra Vol 00:16:14

The mechanism underneath it is the part I care about. How would Anthropic know? They know because they see the prompts. So every one of these misuse reports is two things at once: the best available evidence about what people do with these tools, and evidence produced by a party with an interest in what it shows. Not an accusation — just how the reporting works.

49. Lenar Kess 00:16:36

And Science ran a piece today that makes exactly that discomfort visible. The headline is "'Chilling' warning or overreaction?" — about a recent AI bioweapons report splitting the expert community. The people who study this professionally don't agree on how to read these findings.

50. Damra Vol 00:16:52

That piece is the most useful thing in the segment. When the labs say misuse is escalating, there's a reflex to treat disagreement as denial. The Science reporting shows the disagreement is methodological — about uplift, about baselines, and about whether a model that helps someone move faster has changed what they could actually do. That's a dispute about method among people who do this for a living.

51. Lenar Kess 00:17:16

Meanwhile Cohere's chief executive told CNBC that AI models are becoming the most potent cyber weapon ever created. Which is a strong claim from the same company that spent the weekend asking who gets to define the rules.

52. Damra Vol 00:17:30

Yeah, and I don't think those are inconsistent. You can believe the technology is dangerous and also believe the three largest vendors shouldn't be the ones writing the rulebook. Those beliefs live together fine.

53. Lenar Kess 00:17:42

One line of context on where state capacity is heading, and then we move. The Washington Post, sourced to current and former officials, reports the National Security Agency plans to reorganize into five units — artificial intelligence, China, cybersecurity, warfighting, and global intelligence — under General Joshua Rudd. There's no confirmed structure, no timeline, and no named director for the AI unit. Once one of those exists, it'll tell you more than the reorg rumor does.

54. Lenar Kess 00:18:12

The data center siting fight got a tactic today. Ann Davis Vaughan at The Information reports that Amazon, Microsoft, Oracle and other developers are now siding with consumers on power costs — sweetening financial offers to municipalities and regulators to get facilities approved.

55. Damra Vol 00:18:30

Siding with consumers is a generous phrase for it. Developers are offering to eat some of the rate impact so the local utility commission doesn't have to explain a bill increase to voters. They're buying consent, and I mean that descriptively. It's what you do when the permitting fight is the actual constraint on your build schedule.

56. Lenar Kess 00:18:49

And it's a sign the constraint moved. A year ago the binding limit was chips. Then it was interconnect queues and transformers. Now it's the county board.

57. Damra Vol 00:18:58

On the other side, the Guardian ran an op-ed from Tyler Turner arguing labor unions should organize around a construction moratorium until, in his words, livelihoods and communities are protected. It's an argument, not reporting, and he's making it from a stated position. But his number is one to hear: developers plan roughly eight trillion dollars of build-out, which he puts at about a quarter of US gross domestic product.

58. Lenar Kess 00:19:26

His theory of the bet is direct too — that the spending pencils because AI lets businesses cut their largest expense, which is labor. So the workers building the facilities are, on his account, financing the thing that replaces them.

59. Damra Vol 00:19:40

Whether or not you buy that, it's the first time I've seen someone put the siting fight and the labor fight in the same sentence with a proposed tactic. A moratorium is a bargaining position, and bargaining positions are how permitting fights actually resolve.

60. Lenar Kess 00:19:55

There's also a health claim in the pile. Justine Calma at The Verge has the Environmental Protection Network — a group of former EPA officials — identifying thirty federal actions since January 2025 that they say worsen health risks from data center pollution. That's their count, from a group with a stated view, so take it as an advocacy finding rather than a regulatory one.

61. Damra Vol 00:20:18

Now here's the item that sits oddly next to all of that. Reuters got a document showing Samsung and SK Hynix <emphasis>rejected</emphasis> a proposal from Korea Electric Power to prepay about eighteen-point-seven billion dollars — twenty-five trillion won — of power bills for their chip cluster. Their stated reason: uncertainty over long-term chip demand.

62. Lenar Kess 00:20:41

So one set of companies is paying municipalities for certainty about siting, and another set is refusing to pay a utility for certainty about demand.

63. Damra Vol 00:20:50

I wouldn't push that into a demand-collapse thesis — a twenty-five-trillion-won prepayment is an enormous ask, and refusing it is ordinary negotiation. But the reason they gave is on the record, and it's the same word the market used on Monday: uncertainty about how long this lasts.

64. Lenar Kess 00:21:08

And the capital keeps finding the building side regardless. Marissa Newman at Bloomberg reports Buildots — the US-Israeli startup using computer vision to speed up construction of data centers and chip fabs — raised a hundred and thirty million dollars at a near-billion-dollar valuation. Somebody is very confident the concrete gets poured.

65. Lenar Kess 00:21:28

Three quick ones to close. First, robotics — and I'll say up front that arXiv's robotics section dumped ten papers at an identical four a.m. timestamp today, which is a fact about the announce batch rather than about robotics. Two of them are about what happens once these systems run in production.

66. Damra Vol 00:21:45

    Start with the unpleasant one. It's called DropVLA, from a group at Fudan and City University of Hong Kong. It's an action-level backdoor attack on vision-language-action models — the models that take camera input plus a language instruction and emit robot actions. Poison zero-point-three-one percent of the training episodes, and you can force a chosen action primitive, something like open gripper, to fire at a moment you pick.

67. Lenar Kess 00:22:13

    What's the success rate?

68. Damra Vol 00:22:14

    Between 98.67 and 99.83 percent, with clean-task performance preserved above 98.5. So the robot does its actual job perfectly, and drops whatever it's holding when it sees your trigger. It fires within twenty-five control steps, about fifty milliseconds at 500 hertz. They validated it on a seven-degree-of-freedom Franka arm, not just in simulation. And vision triggers work while text triggers mostly don't — text-only transfer was zero-point-seven-two percent.

69. Lenar Kess 00:22:48

    It's a v5 preprint with self-reported numbers, and the attack assumes you can poison the fine-tuning data, which is a large assumption.

70. Damra Vol 00:22:56

    Right. But pair it with the other paper. Robion, from a group including Christina Giannoula, is a serving system for exactly these models. It splits the vision-language stage and the action-diffusion stage across two streams <emphasis>within</emphasis> a single graphics processor, because the stages are millisecond-scale and disaggregating them across separate chips doesn't work. They report serving up to sixty-four robots from one four-GPU edge server while hitting ninety-eight percent of their service-level objectives, and six-point-seven times the robot load of vLLM-Omni.

71. Lenar Kess 00:23:31

    Sixty-four robots sharing four chips and a fine-tuning pipeline.

72. Damra Vol 00:23:36

That's the deployment that actually exists today. And there's a third preprint in the batch, GuardrailLoop, with a finding I keep returning to. Across two hundred forty injected crashes, every run recovered the correct outcome, but only two hundred ten preserved the execution trace. Thirty pre-commit crashes silently repeated a planner call. Recovering the answer isn't evidence you ran once.

73. Lenar Kess 00:24:01

Second item, and it's short. Elon Musk posted that Grok 4.8 is a two-point-five-trillion-parameter model trained on xAI's new C++ software stack. He says training finishes this week and the model moves to reinforcement learning after. He also says Grok Bot now routes its network traffic through your local machine, and he spent Sunday promoting a live stream of someone building a company with it.

74. Damra Vol 00:24:26

The local-egress change is the more concrete claim of the two. If the agent's requests leave from your machine rather than xAI's, it inherits your network position: your logins, your internal hosts, your rate limits, and your address reputation. That changes what it can reach without changing anything about the model.

75. Lenar Kess 00:24:46

And the timing on the training claim is its own comment, given he co-signed the pacing call the same weekend. I'm not going to narrate that as hypocrisy — self-reported ship dates from that account slip constantly, and this isn't a release. It's a date he put on the record.

76. Damra Vol 00:25:02

Last one. OpenAI published a short video with a co-founder of Fyxer, an AI executive assistant. His claim: zero to thirty million dollars in annual recurring revenue in the first year, built entirely on OpenAI models, with no proprietary training at all. They compressed what he describes as five hundred thousand hours of executive-assistant workflow into months of development.

77. Lenar Kess 00:25:27

That's a vendor-published video and the revenue figure is a customer's self-report inside it, so weight it accordingly. What's the architecture?

78. Damra Vol 00:25:35

It digests email, extracts intent, pulls historical context, and reranks so the urgent thing surfaces first. Draft generation is several OpenAI components chained together with a feedback loop tuned to cut token cost. Asked what the moat is, he doesn't say inference quality — he says workflow mapping and prompt design. And he mentions the team has a Slack channel straight to OpenAI engineers, which is its own kind of advantage.

79. Lenar Kess 00:26:05

    Ethan Mollick made the mirror-image point over the weekend, from the sovereignty side. Organizations that want their own maintained small models keep discovering how expensive it is to hold a fine-tuned model ahead of the general-purpose one that improves under them every quarter.

80. Damra Vol 00:26:21

    Same finding from two directions. For most people right now the advantage comes from knowing what the work actually is, not from owning weights. Which is also why IEEE Spectrum's Eliza Strickland profile of Andon Labs is fun without being a business result. They hand agents actual businesses and let it get weird on purpose. Those are testbeds, and the company says so.

81. Lenar Kess 00:26:45

    Two things on the calendar out of today. Jeffries has his caucus in a room tomorrow trying to decide what House Democrats want, with proposals ranging from a subpoena-powered select committee to twenty-year prison sentences. And Musk says that training run finishes this week. Both of those have dates attached, which is more than the pacing pledge managed in three days of unanimous agreement.

82. Damra Vol 00:27:07

    The lease is what settles the rest of it. Six years, thirteen-point-seven billion dollars, a Georgia site that hasn't been poured. If the frontier really does get paced, somebody ends up sitting on capacity they described to investors as necessary.

Hosts on this episode

- Lenar Kess moderator
- Damra Vol critic