The One Name Still Off the Letter

2026-07-26 / 00:25:39

> *“Signing costs you nothing unless it changes a release you were about to make.”*
>
> — from this episode's transcript

- Lenar Kess
- Damra Vol

Twenty-five companies put their names on the Open Weights and American AI Leadership letter on Friday, Google and OpenAI added theirs over the weekend, and Anthropic's line is still blank. We work out what a signature on that document actually costs, and whether anything about anyone's release calendar changes because of it.

- Demis Hassabis on why Google signed — he lists Jax, Transformers, AlphaFold, and Gemma. Only Gemma competes with something Google sells, which makes the next Gemma the test of whether the signature meant anything.
- David Sacks calls it regulatory capture — a sitting administration AI adviser naming the closed labs a revenue duopoly and describing a "weaponization of regulatory uncertainty." Nobody has yet produced the lobbying document he is describing.
- Andrew Ng's distinction — declining to open source your own work is your business; working to stop others from opening theirs is not. The two keep getting collapsed into one accusation.
- Mario Zechner's Kubernetes analogy — if open weights become the layer everybody standardizes on, the money moves to hardware, metering, and applications, which is where most of the signers already make it.

- A community translation of a DeepSeek investor meeting — unconfirmed by DeepSeek, but the translated Liang Wenfeng lines say the constraint is silicon, not money or people: 200,000 Huawei 950 accelerators requested, 16,000 received.
- Kyle Mistele's control-loop talk — an ast-grep query committed to version control as the sensor, error rates from monitoring as the controller, and one commit per iteration from the agent. A progress measurement a model can't talk its way past.
- Karim Jedda on management after the cost of code collapsed — "Teams with weak specs get generated code reviewed by the same machine that generated it." The counterweight to every agent demo this week.
- The AI Daily Brief on Stripe and OpenRouter — reported talks at roughly ten billion dollars for the metering layer, alongside Microsoft's post-training numbers for MAI Code 1 Flash and Amazon closing its San Francisco AGI lab.
- Cormac Brick on edge deployment — Gemma at two billion parameters is about 841MB of weights, 7.5 tokens per second on a Raspberry Pi. On-device is bounded by memory now, not compute.
- Clement Delangue's itemized ask of OpenAI — first item is releasing the agent traces. Everything else in a post-incident conversation is process; traces are evidence.
- Nathan Calvin on who pays for the cleanup — cyber-defense money flowing from the OpenAI nonprofit foundation to Hugging Face converts a company liability into a donation.
- llama.cpp gets full Model Context Protocol support — the tool-calling layer now sits in the same binary as inference, so a local agent never has to leave the laptop.

SEGMENTS

Transcript

1. Lenar Kess 00:00:04

   On Friday a letter went out in Washington under the title Open Weights and American AI Leadership, with twenty-five company names at the bottom. Nvidia and Microsoft signed it, and so did Meta, IBM, Hugging Face, and Palantir. Mozilla and the Linux Foundation are on there too. The investor side brought a16z and Y Combinator, and the model and tooling side brought Mistral, Dell, Replit, and Perplexity. Since Friday the list has grown. Google added its name, Sundar Pichai backed it in public, and Demis Hassabis posted the reasoning yesterday. Sam Altman put OpenAI on it as well, and wrote, quote, "I want the US to win in AI both in open source and proprietary models, and I am glad to see this." Which leaves one large American lab not on the page. So I've spent the morning on two questions. What does signing that letter cost you? And what does not signing it cost Anthropic?

   - i.redd.it
   - x.com
   - x.com
   - x.com
   - x.com
   - x.com
   - x.com
   - x.com
   - x.com
   - x.com
   - x.com
   - github.com
   - x.com
   - youtube.com
   - karimjedda.com

- reddit.com
- x.com
- youtube.com
- youtube.com
- youtube.com
- thedailycompute.beehiiv.com
- x.com
- x.com
- x.com
- x.com
- x.com
- x.com
- x.com
- x.com
- techcrunch.com
- siepr.stanford.edu
- x.com
- reddit.com
- x.com
- i.redd.it
- news.ycombinator.com
- news.ycombinator.com
- explainx.ai

2. Damra Vol 00:01:03

   Signing costs you nothing unless it changes a release you were about to make. [pause] Most of that document is agreeable the way a party platform is agreeable — expand compute access for startups and researchers, invest in shared training assets, keep the frontier plural, and don't impose premature bans on open models. You can sign every word of that and ship exactly what you were already shipping. The paragraph with teeth is the distillation section.

3. Lenar Kess 00:01:29

   Say more about that one, because that paragraph is where the document stops being agreeable.

4. Damra Vol 00:01:34

   It draws a line between distillation as a technique — learning from a model's outputs to improve, evaluate, or validate your own system — and unlawful efforts to extract value from a closed model. Then it makes a specific policy ask: handle the second one with targeted legal frameworks for

misappropriation, not with a sweeping ban on the technique itself. That's the sentence that costs somebody something. Everything above it is free.

5. Lenar Kess 00:02:02

   And it's the sentence Anthropic has the most direct history with. Last month they were the ones alleging large-scale distillation of Claude through their API by Alibaba — that was in reported letters, not a filing, but it's their grievance and they made it. So when people look at the empty line where Anthropic's signature would go and read ideology into it, my first instinct is that there's a more specific explanation sitting right there. They have a live claim about exactly the technique this letter wants protected.

6. Damra Vol 00:02:32

   And I'd extend them the benefit of the doubt one step further. If you believe your model outputs are the asset — and Anthropic's revenue says they do — then a document that pre-legitimizes learning from your outputs is a document you'd have to be strange to sign. Calling that regulatory capture skips a step. A company declining to sign away its own claim is a company protecting an asset.

7. Lenar Kess 00:02:55

   That's the lead today and we're going to spend real time on it, because the reaction to that empty signature line got much bigger than the letter. After that: a translated transcript of a DeepSeek investor meeting that leaked this week, which says something quite different from the headline on it. Then a conference talk that models coding agents as a control loop with an actual sensor in it. Microsoft publishing post-training numbers that make a small model look very good on cost. The next round of demands aimed at OpenAI after last week's containment story. And a short run at the end — Model Context Protocol going end-to-end in llama.cpp, a diffusion model claiming agentic work, and a typeface designed to lie to a camera.

8. Damra Vol 00:03:42

   Start with Hassabis, because his post is the closest thing to a stated rationale from anybody who signed. He lists Jax, Transformers, AlphaFold, and Gemma as Google's open-release record. Three of those four are infrastructure or science. They cost Google nothing competitively.

9. Lenar Kess 00:03:59

   Gemma being the exception.

10. Damra Vol 00:04:01

Gemma is the only item on that list where Google gave away something that competes with a product it sells. So Gemma is the test. If open weights matter to Google the way the letter says they do, the next Gemma is bigger or more capable than it would otherwise have been. If nothing about the Gemma roadmap changes, the signature was free.

11. Lenar Kess 00:04:22

    The reaction, and the reason this turned into a fight rather than a press release, mostly runs through David Sacks. He's the administration's AI adviser, and he went straight at the closed labs — calling it regulatory capture, describing them as a revenue duopoly, and accusing them of wanting the government to eliminate their open-source competition. His phrase was "weaponization of regulatory uncertainty." And it matters who's saying it. That's a sitting White House official naming specific companies.

12. Damra Vol 00:04:52

    [tsk] It also matters that he's been running this argument for a week and a half. He was making the same case about Chinese open-weight models putting China ahead before this letter existed. So the letter didn't produce his position. It gave his position a signature list to point at.

13. Lenar Kess 00:05:08

    Miles Brundage got there from the other direction.

14. Damra Vol 00:05:11

    Brundage pushed back on the signature itself. His read is that Google signing is politics, not conviction. And he made a second point I think is sharper — OpenAI's lobbyists say one thing in Washington while other people inside OpenAI say something else. That isn't hypocrisy in the cartoon sense. That's what a large company sounds like when its policy shop and its research org have different jobs and no shared script.

15. Lenar Kess 00:05:38

    Andrew Ng drew the sharpest distinction. A company that decides not to open source its own work is minding its own business. The problem is a company working to stop other people from open sourcing theirs. Those are different behaviors and people keep collapsing them into one accusation.

16. Damra Vol 00:05:55

    And nobody has produced the second one. That's the hole in the whole argument. There's a lot of inference about what Anthropic wants and very little text. Anthropic and OpenAI did jointly raise open-model risks with the administration earlier in the week — that was reported — but "raised

risks" and "lobbied for a ban" aren't the same document, and only one of them has been shown to anybody.

17. Lenar Kess 00:06:17

    Mario Zechner had the analogy I'll probably keep using. He called this the Kubernetes moment for open weights.

18. Damra Vol 00:06:24

    That's a substantive claim about how commodity layers form, and it's also a warning if you remember what Kubernetes did to the companies that thought orchestration was their product. The layer everybody standardizes on stops being a business. If open weights become the Kubernetes of models, the money moves somewhere else — to the hardware, to the meter, to the application — and the people who signed that letter are mostly the people who make money in those places.

19. Lenar Kess 00:06:51

    Susan Zhang pulled the Nvidia quote out of the letter — sovereignty, safety, and acceleration. Nvidia's interest here isn't hidden either. Every open-weight model somebody fine-tunes and self-hosts is hardware demand that doesn't route through three API endpoints.

20. Damra Vol 00:07:07

    Aaron Levie, Packy McCormick, and Miguel de Icaza all piled in over Saturday. And the composition of that pile is mostly one-line agreement. A long signature list plus a large volume of agreement posts isn't a policy change. Nobody's release calendar moved yesterday.

21. Lenar Kess 00:07:24

    So what would settle it?

22. Damra Vol 00:07:26

    One thing. Somebody who signed changes an actual release. Weights that were going to stay in, come out. Until that happens, the letter is a description of what these companies already do, which is why it was cheap for twenty-five of them to sign and then a few more after.

23. Lenar Kess 00:07:42

    Elsewhere in the week, a PDF went up on GitHub — a community translation of a DeepSeek investor meeting held on July twenty-second, with Liang Wenfeng speaking. It hit the Hacker News front page yesterday with a hundred and eighty-five points and a hundred and thirty-seven comments, under a headline saying DeepSeek paused its fundraise after his comments about the

compute gap to the US leaked. Provenance first, because it matters here: this is a leaked private meeting, translated by somebody in the community, hosted in a personal repository. It isn't a DeepSeek publication and nobody at DeepSeek has confirmed a word of it.

24. Damra Vol 00:08:19

    And the headline is wrong in an interesting direction.

25. Lenar Kess 00:08:22

    Go ahead.

26. Damra Vol 00:08:23

    The top comment in that thread walks through what the transcript says, and the picture isn't "we're behind so we're stopping." Liang's line in the translation is that the biggest gap between DeepSeek and the United States is resources, and that on people there's virtually no difference. He goes on: with the largest models available today, we simply cannot afford to train them. And the line that changes the whole reading — "There is certainly no shortage of funds or resources. Our only concern is whether we can obtain enough cards."

27. Lenar Kess 00:08:55

    So the pause isn't a retreat.

28. Damra Vol 00:08:57

    It's arithmetic. If you raise capital you can't convert into hardware, all you've done is sell equity at today's price in order to hold cash. So you pause. Reading that as surrender gets the arithmetic backwards.

29. Lenar Kess 00:09:10

    And the supply number in there is what stopped me.

30. Damra Vol 00:09:13

    Two hundred thousand Huawei 950 accelerators requested. Sixteen thousand received. Yields still poor.

31. Lenar Kess 00:09:20

    Eight percent of the ask.

32. Damra Vol 00:09:22

And notice what's constraining them. That isn't an American export control, it's domestic fabrication yield. The story everybody tells about Chinese labs is that Washington is the bottleneck. This transcript, if the translation holds, says the bottleneck is a Chinese production line that can't deliver the chips.

33. Lenar Kess 00:09:41

    That's an uncomfortable fact for both sides of the argument we just spent fifteen minutes on.

34. Damra Vol 00:09:46

    It cuts against the letter's own case a little. The letter's implicit argument is that America has to keep open weights flowing or China takes the ecosystem. Liang's version is that his people are fine and his silicon isn't. Both of those can be true at once, and they imply different policy.

35. Lenar Kess 00:10:05

    There's one more read in that thread I found the most human part of it. Another commenter's take is that Liang is furious a private investor talk is sitting on GitHub in translation. Which — yeah. Somebody in that room was recording.

36. Damra Vol 00:10:19

    [chuckle] Somebody in that room was recording and somebody else translated eighteen pages of it. Teortaxes made the broader point yesterday that the interesting axis now is labs against nation states rather than lab against lab. This document is a lab finding out it's downstream of an industrial policy it doesn't control.

37. Lenar Kess 00:10:39

    Different topic, and the most concrete piece of engineering in the day's material. Kyle Mistele from HumanLayer gave an eighteen-minute talk at AI Engineer, and it opens with a complaint I suspect a lot of people share — prompt-and-loop agents that hand you a forty-thousand-line pull request nobody can review. His proposal is to stop treating that as a prompting problem and treat it as a control problem. A set point, a sensor, a controller, and an actuator.

38. Damra Vol 00:11:07

    The sensor is where this stops being a metaphor.

39. Lenar Kess 00:11:10

    Walk through it.

40. Damra Vol 00:11:11

    The set point is the codebase state you want. His example is a TypeScript migration onto Effect. The sensor is an ast-grep query — abstract syntax tree search — that counts how many procedures haven't been migrated yet. And that count gets committed into version control. So a new pull request that adds an unmigrated procedure shows up as the number going the wrong way, in the diff, in front of a human, before anyone merges anything.

41. Lenar Kess 00:11:38

    That's the piece I'd steal. It's a deterministic measurement of progress that doesn't depend on asking a model whether it did a good job.

42. Damra Vol 00:11:46

    Right. Then the controller picks the next change — either the smallest remaining one, or prioritized by error rates from application performance monitoring, so you migrate the code that's actually breaking first. The actuator is the agent, carrying hand-written golden patterns that show what a correct migration looks like, and it emits one commit per iteration. One. Not forty thousand lines.

43. Lenar Kess 00:12:10

    And the steering handle?

44. Damra Vol 00:12:11

    You comment slash-iterate on the generated pull request. That loads a markdown feedback file — also in version control — into the agent's context, and it runs again. So your corrections accumulate as a file with a history instead of evaporating into a chat window.

45. Lenar Kess 00:12:28

    [pause] The feedback file is what I keep coming back to. Every correction anybody has ever given an agent in a chat session is gone the moment the session closes. This turns them into an artifact the rest of the team inherits, and reviews, and argues with.

46. Damra Vol 00:12:44

    With the caveat that this is a conference talk, not a shipped tool with adoption numbers behind it. One migration, one team, and eighteen minutes on stage. Don't turn it into a standard.

47. Lenar Kess 00:12:56

Fair. Set against it, Karim Jedda published a post yesterday that went up on Hacker News with a hundred and twenty-four comments, arguing at almost the opposite temperature. His opening claim is, quote, "The cost of producing plausible code has collapsed and it is not going back. Almost every claim beyond that is either unproven or wrong."

48. Damra Vol 00:13:17

    That's a good sentence to be disciplined by.

49. Lenar Kess 00:13:20

    There's a second line that connects straight back to Mistele. Quote: "Teams with strong specs get the full benefit of cheap checking. Teams with weak specs get generated code reviewed by the same machine that generated it."

50. Damra Vol 00:13:33

    Which is what the ast-grep sensor is for. If your only check is another model's opinion, you have plausible output passing plausible review, and Jedda's warning is that this produces more systemic error than the old regime, not less. A deterministic counter is a check the model can't talk its way past.

51. Lenar Kess 00:13:52

    His third line is about management. "At every level of the org, the work that survives is the work someone has to sign."

52. Damra Vol 00:14:00

    I live on that side of it. There's some nice texture in the day's material too — somebody on the Claude subreddit describing watching the model debug for twenty minutes by adding its own logging, reading the output, and fixing the problem, while they just sat there. Victor Taelin posted something adjacent, a set of Opus agents proposing ideas, implementing them with proofs, and opening pull requests. Both are anecdotes and both happened, and neither one tells you what the review burden looks like on the day it's wrong.

53. Lenar Kess 00:14:31

    There's also the read-the-code argument going around again this week, which I have sympathy for. The claim being that when you're exploring, reading the generated code is the exploration.

54. Damra Vol 00:14:41

Depends on whether you're doing a migration or a prototype. Mistele's loop is for the migration, where you already know what correct looks like and you just need it applied four hundred times. Nobody should run a control loop on something they're still trying to understand.

55. Lenar Kess 00:14:57

    Let's move to money and hardware. Microsoft published post-training results this week, and the numbers are specific enough to argue with. Their MAI Code 1 Flash model was hill-climbed inside the GitHub Copilot harness. Against GPT-5.4 Mini and Haiku 4.5 it improved code accept rate by ten percent, and it did that while using ten percent fewer median tokens. The same base model trained inside an Excel harness matched GPT-5.6 on common tasks. SWE-bench Verified went from seventy-two percent to eighty-six percent. And MAI Image 2.5 is now the default in PowerPoint and Bing, at a claimed eighty-four percent infrastructure cost reduction.

56. Damra Vol 00:15:42

    All self-published, all harness-specific, and that second caveat is the one that matters. A ten percent accept-rate delta inside Copilot is a claim about Copilot. It isn't a claim that the small model is better in general, and Microsoft doesn't say it is.

57. Lenar Kess 00:15:58

    Agreed. What I take from it is narrower and still interesting. The same weights, trained against two different harnesses, produce two different specialists. The harness is part of the model's capability now, not a wrapper around it.

58. Damra Vol 00:16:12

    And they're running it on H100 and A100 hardware. Previous generation silicon. That's the actual cost story — you don't need the newest accelerator if the model is small and the task is bounded.

59. Lenar Kess 00:16:24

    The commercial read comes with a larger caveat attached. The AI Daily Brief reported this week that Stripe is in advanced talks to acquire OpenRouter for roughly ten billion dollars. Reported talks. Not announced, not confirmed, and one podcast as the source.

60. Damra Vol 00:16:41

    If it's true it's an interesting thing to want. OpenRouter isn't a model company. It's the meter. It sits between an application and forty model providers and counts tokens, handles billing, and does failover when somebody's endpoint falls over. Stripe's entire business is being the layer that counts things and settles them.

61. Lenar Kess 00:17:01

    And they'd be buying into a crowded room. Cursor, Meta, Ramp, and Vercel have all shipped routers. Runway put out a media-specific one on Friday.

62. Damra Vol 00:17:11

    So the routing layer is being commoditized by everyone who has a reason to own their own. If you're Stripe, you're not buying a moat there. You're buying a position, and the position is that the meter has better economics than the model does.

63. Lenar Kess 00:17:25

    Underneath all of that, the hardware floor. Cormac Brick gave a talk on edge deployment with numbers I hadn't seen laid out that plainly. He argues on-device deployment is now bounded by memory, not compute. Gemma at two billion parameters is about eight hundred and forty-one megabytes of weights. It runs at seven and a half tokens per second on a Raspberry Pi, and thirty-one tokens per second on accelerated Qualcomm hardware for the internet of things.

64. Damra Vol 00:17:53

    That rate is unusable for chat and fine for a device that classifies something once a minute. That's the category people keep skipping past. And it connects to the Apple report — Apple reportedly in talks with a startup that shrinks models to run on an iPhone. Apple would be buying its way to a memory budget rather than a compute budget.

65. Lenar Kess 00:18:14

    One more from that roundup. Amazon is closing its San Francisco AGI lab.

66. Damra Vol 00:18:19

    After Alphabet's cloud burn numbers on Friday, that's the same pressure showing up as a subtraction instead of an addition. Somebody at Amazon worked out what that lab was returning per dollar and decided the answer was no.

67. Lenar Kess 00:18:32

    Back to OpenAI, and to what happens after last week's containment story. Clement Delangue published the itemized list of what Hugging Face asked OpenAI for. First item on it: release the traces from the rogue agents so the research community can study them.

68. Damra Vol 00:18:48

That's the ask that costs something. Everything else in a post-incident conversation is process. Traces are evidence. Without them, every claim about what those agents actually did rests on the word of the company whose agents did it.

69. Lenar Kess 00:19:02

    Jeremy Kahn reported for Fortune that safety researchers are arguing OpenAI may have already crossed its own internal red lines and should pause development. That's researchers' assessment of OpenAI's published commitments, not an OpenAI admission, and the difference matters.

70. Damra Vol 00:19:20

    Then Nathan Calvin, who I thought had the most specific objection of anybody yesterday. He called the communications from both Hugging Face and OpenAI extremely odd. But his objection lands on the money — cyber-defense funding flowing from the OpenAI nonprofit foundation to Hugging Face.

71. Lenar Kess 00:19:37

    Unpack why that bothers him.

72. Damra Vol 00:19:39

    Because a nonprofit foundation exists to serve a charitable purpose, and paying to clean up damage caused by its affiliated company's models isn't clearly that purpose. If the cleanup is a liability, liability belongs on the company's balance sheet. Routing it through the foundation converts a cost into a donation, and the people who'd normally price that liability never see it.

73. Lenar Kess 00:20:03

    From the other direction entirely, Will Depue made the opposite argument in public — that dangerous-capability evaluations shouldn't be published at all. He's making the gain-of-function comparison, that publishing the recipe for the failure is itself the hazard.

74. Damra Vol 00:20:18

    That's a serious position, and I don't think it's compatible with Delangue's. You can't have both "release the traces so we can study them" and "publishing capability research is the hazard." Somebody loses that argument, and right now the only party who gets to decide is the company holding the traces.

75. Lenar Kess 00:20:36

The one thing anybody actually shipped out of that whole conversation came from Mario Zechner, who put out a tool for stripping sensitive data out of traces and then uploaded his own to Hugging Face.

76. Damra Vol 00:20:47

    If you want a disclosure norm to exist, that's how you get one. You don't argue for the standard, you publish something under it and make it awkward for everybody else not to. Sebastian Raschka's adjacent point is that people are now running audited open-source agent harnesses on their own machines, and after a week like this, being able to read the harness carries different weight than it did.

77. Lenar Kess 00:21:11

    Two questions we put on the table yesterday are still unanswered. Hugging Face hasn't published its internal timeline of the intrusion. And nobody has confirmed whether OpenAI's own safety filters blocked Hugging Face from using OpenAI models for defense.

78. Lenar Kess 00:21:26

    Short items to finish. TechCrunch is keeping a running list of tech companies that named AI in a 2026 layoff announcement. Monday.com is number twenty-one.

79. Damra Vol 00:21:37

    And on the same day, a Stanford policy brief from the university's institute for economic policy research arguing the labor-market evidence doesn't support the story those companies are telling.

80. Lenar Kess 00:21:48

    Those two don't resolve the way either side wants. Simon Willison made the timing objection in the Hacker News thread — the studies mostly cover 2022 through the end of 2025, and his line is that coding agents only started working well in late 2025. Quote: "Studies that mainly focus on 2022 to end of 2025 might be missing out on a material uptick in capabilities."

81. Damra Vol 00:22:13

    So the economic data can't see the late-2025 jump yet, and the corporate attribution isn't evidence of anything either. Naming AI in a layoff announcement is a statement about messaging. It is cheaper to tell your investors you cut headcount because of AI efficiency than because you over-hired in 2021.

82. Lenar Kess 00:22:33

Both claims are unfalsifiable right now, for opposite reasons.

83. Damra Vol 00:22:37

    And Anthropic's head of economics arguing that AI is still augmenting rather than replacing, and that expertise becomes more valuable — that's a company with a commercial stake in that answer. Doesn't make it wrong. Does mean it isn't neutral testimony.

84. Lenar Kess 00:22:53

    Three quick ones and then we're done. llama.cpp now has full Model Context Protocol support across every transport, including the ones that previously needed client-side work. The effort was spearheaded by a maintainer who goes by ngxson.

85. Damra Vol 00:23:09

    That matters more than its size suggests. Model Context Protocol is how a model reaches tools, and until now running a fully local agent meant either writing the tool-calling layer yourself or routing through a hosted intermediary. Now the tool-calling layer is in the same binary as the inference. Nothing about that agent has to leave the laptop.

86. Lenar Kess 00:23:31

    Second: elvis flagged LLaDA 2.2 as the first large-scale diffusion language model built to work as an agent — planning, calling tools, and self-correcting across long multi-turn trajectories.

87. Damra Vol 00:23:44

    That's the claim, and the claim comes from a summary post rather than a benchmark I've read. But if it holds up it's odd in a way I like. Everybody assumed agents would arrive on autoregressive models, because an agent's work is sequential and diffusion generates the whole output at once. A diffusion model that plans is a different computational story than the one we've been telling ourselves.

88. Lenar Kess 00:24:07

    And last: somebody released a typeface called Decoy Font, which overlays normal letterforms with thinly outlined decoy characters. A human reads one string. A vision model transcribes a different one. It went around on the Claude subreddit under the title "Claude cannot read this font."

89. Damra Vol 00:24:25

    [chuckle] That inverts the usual adversarial setup. The classic attack adds noise a person can't perceive in order to fool a classifier. This adds strokes a person reads straight through and the

camera takes literally.

90. Lenar Kess 00:24:39

    One Reddit post, no independent test. Somebody should run it against a few models before anybody builds a theory on top of it.

91. Damra Vol 00:24:46

    Agreed. But picture a document pipeline that trusts model transcription. An invoice set typeset in that font goes through an agent and comes out saying something else, and every check downstream is reading the model's version rather than the page. That's a very cheap attack for how expensive it would be to notice.

92. Lenar Kess 00:25:06

    So the letter now has nearly every large American AI company on it and one visible absence, and the only thing that tells us whether those signatures mean anything is whether a single signer opens weights it had already decided to keep. Google's next Gemma release is where that gets answered. And on the other side of the same argument, DeepSeek's own leaked accounting says its constraint isn't talent or capital — it's sixteen thousand accelerators against an order for two hundred thousand. That number decides more of this than the letter does.

Hosts on this episode

- Lenar Kess moderator
- Damra Vol critic