Pray for Rain, Approve the Datacenter

2026-06-08 / 00:25:31

> *“We spent a decade teaching people not to pipe curl into bash. The agent config file is the new version of that, except it fires the moment you open the repo.”*
>
> — from this episode’s transcript

- Lenar Kess
- Damra Vol

The consequential part of an AI system keeps moving out of the model and into the wrapper around it — the cooling loop, the org chart, the config file, the ownership structure — and the tools we use to trust that wrapper are running behind it. Five stories, one recurring tension.

- The Guardian finds about two-thirds of 809 planned US datacenters are slated for drought-hit land; closed-loop cooling saves water but trades it for fossil power that needs water of its own.
- OpenAI’s enterprise talks feature banks rebuilding their orgs: Allica Bank collapsing roles into ”squadlets,” Erste Group budgeting for a full platform rewrite every 18 months, plus ChatGPT-in-Excel Skills and Codex — held against one engineer’s MCP catalog server.
- The Miasma worm: one dropper wired into seven config files across Claude Code, Gemini, Cursor, VS Code, npm, Composer, and Bundler — opening a cloned repo becomes an execution event.

- Schneier and Nathan Sanders argue against Bernie Sanders' equity-stake plan, proposing energy taxes and an AI Public Option instead — set against Korea's GPU program and NVIDIA's UK sovereign-AI post.
- Two arXiv papers on measuring safety too late: Attack Selection shows strategic timing drops measured control safety 20-28 points, and Don't Just Fix It in Post argues the science belongs in training dynamics, not the finished snapshot.

SEGMENTS

Transcript

1. Lenar Kess 00:00:04

   Last year the governor of Utah asked residents to pray for rain. The Great Salt Lake was shrinking, the state was deep in drought, and the official ask was prayer. Last month, that same state approved one of the largest datacenters in the world — a complex twice the size of Manhattan — in a county that's been in drought since the summer before. So here's the question I keep turning over: when there isn't enough water to go around, who gets cut first, the residents or the servers?

   - theguardian.com
   - theguardian.com
   - axios.com
   - youtube.com

- youtube.com
- youtube.com
- youtube.com
- reddit.com
- safedep.io
- msit.go.kr
- msit.go.kr
- blogs.nvidia.com
- arxiv.org
- arxiv.org

2. Damra Vol 00:00:30

   And that tension — who gets cut when the resource runs short — is a decent map for the whole day. We've got the water fight under the datacenters; banks on stage describing how they rebuilt their engineering orgs around AI agents; a worm that turns opening a repo into an attack; a fight over whether the government should own the AI labs or tax them; a couple of national chip-buying programs; and two papers arguing we're measuring safety wrong. Different stories, one recurring question: as the capability moves into the system around the model, who controls it, and can you trust it? Start with the water.

3. Lenar Kess 00:01:06

   The Guardian ran the numbers this morning — Oliver Milman, with Andrew Witherspoon on the data side. They looked at 809 planned US datacenters. 517 of them, about two-thirds, are slated for places that have been in drought for the past year. Roughly the same proportion holds for the ones already running. And more than 60% of the contiguous US is in some stage of drought right now, the largest spring expanse in the modern record.

4. Damra Vol 00:01:32

   I want to pin one number down, because the cooling story has a twist. The Guardian says a large datacenter can pull up to five million gallons of water a day for cooling — that's the daily water use of about fifty thousand people, into a single building. But that's the evaporative-cooling figure. The industry's answer is closed-loop cooling, where you recirculate the same coolant instead of boiling off fresh water. That cuts water a lot. Here's the catch the article names directly: closed-loop costs you more energy, that energy mostly comes from fossil plants, and fossil generation needs huge amounts of water itself. So you can move the water draw. You don't necessarily make it vanish.

5. Lenar Kess 00:02:15

Meta's Hyperion project in Louisiana shows exactly that trade. Closed-loop cooling, so they get to say they're being efficient — and the facility needs the power of about ten gas-fired plants, which burn fuel and use their own water. Meta told the Guardian the site could use as much as a billion gallons a year, drawn from an aquifer that farmers currently rely on. Their framing is that it'll still use less than if the land were farmed.

6. Damra Vol 00:02:39

   I have to put the counterargument in, because the industry does. The Data Center Coalition's line is that datacenters take a fraction of what agriculture does — even lawn and golf-course irrigation outdraws them. And there's a study from a water company, Xylem, that the Guardian cites: datacenters themselves are only about 4% of the extra water AI will need globally by mid-century. Power generation and chip fabrication eat far more. So if you're only watching the datacenter meter, you're watching the small dial.

7. Lenar Kess 00:03:10

   That's a fair check, and I think both things hold at once. The aggregate share is small; the local concentration is brutal. A datacenter doesn't spread its draw across the country — it sits on one aquifer, in one county, next to specific ranchers. The Guardian quotes Andrew Coppin, who runs a company called Ranchbot that helps ranchers track water use. His line stuck with me: ranchers get told to conserve, not waste a drop, and then a new competitor shows up with near-unlimited access. He said most people, forced to choose, would rather have a beef steak than ChatGPT.

8. Damra Vol 00:03:44

   [chuckle] Hard to argue with the steak. And the trajectory is steep. The same reporting has US datacenters demanding as much as 73 billion gallons of water a year by 2028, up from about 17 billion in 2023. Researchers estimate each hundred-word AI prompt works out to roughly one 500-milliliter bottle of water once you count the cooling. And politically this has stopped being a coastal-environmentalist story — the opposition's coming from rural conservative counties. Polling in the piece says 70% of Americans don't want to live next to a datacenter. New York lawmakers are floating an outright moratorium.

9. Lenar Kess 00:04:25

   What I take from it is that water is becoming a siting constraint with the same teeth as power. For a while the only gating question for a new site was whether you could get enough electricity to it. Now there's a second meter, and in two-thirds of these locations it's already running low. The Utah complex, Stratos, is facing a referendum and a lawsuit from local residents. That's the new cost of doing business in a dry county.

10. Lenar Kess 00:04:50

    Let's move from the water bill to the org chart. There was a cluster of talks out of OpenAI's enterprise event today, and the interesting parts weren't the product demos — they were banks describing how they tore up their engineering structure to absorb this stuff. And I'll flag the source plainly: these are talks at OpenAI's own customer event, so they're part marketing. The operational detail in them is specific enough to take seriously anyway.

11. Damra Vol 00:05:13

    Right — treat them as testimony from a customer the vendor invited on stage. With that caveat, the Allica Bank one is where I'd start. Ravneet Shah, their chief technology officer. Allica is a UK challenger bank for small and medium businesses, licensed in 2019. He says they took AI adoption from 25% of the org to a median workday utilization of 77%, and ran something like 3,700 deployments last year.

12. Lenar Kess 00:05:41

    And the change that's actually a decision, not a metric: they threw out the Spotify squad model and rebuilt around what he calls squadlets. Smaller pods, and they collapsed roles — backend, frontend, and quality-assurance jobs merged into one engineering role; product owner and analyst merged into another. They're trying to grow what he calls product engineers, people who own both the product call and the code. The target is for designers and product staff to ship to production directly by year end.

13. Damra Vol 00:06:14

    That's the claim that makes me sit up, because it's a real bet on where the bottleneck moved. The argument is: when an agent writes a lot of the code, the handoffs between specialists become the slow step, not the writing. So you merge the specialists. For some of the work, that's the right diagnosis. What I'd want to know — and a stage talk will never tell you — is what breaks. When quality assurance is no longer a separate person, who catches the regression the agent introduced with total confidence? In a bank regulated by the Financial Conduct Authority, that isn't rhetorical.

14. Lenar Kess 00:06:51

    He gives a partial answer on the lending side. For underwriting they run what he describes as a hybrid of deterministic and non-deterministic agents — the agent parses broker emails, pulls out the missing fields, auto-fills the application portal, and they've got decision times down under seven to twelve minutes for some applications. The deterministic wrapper is doing the part you can't let the model improvise.

15. Damra Vol 00:07:13

Which rhymes with what the show kept hitting over the weekend — the model is never the whole system, the repair logic around it is. Same pattern here, just with compliance stakes attached.

16. Lenar Kess 00:07:24

    The Erste Group talk added what everyone usually skips. Maurizio Poletto, their chief platform officer. Erste runs a digital banking platform called George. And he's blunt about it: they're already on version two of the platform, they threw away version one, and he expects to throw away version two in about eighteen months.

17. Damra Vol 00:07:43

    That's the detail I trust most in the whole set, precisely because it isn't flattering. Most platform pitches imply you build it once. He's budgeting for a full rewrite every couple of years because the ground under the architecture keeps shifting. He also pushed back on the conversational-everything pitch — their research says retail users prefer button-driven interfaces, they don't show up with prompts ready. So Erste mixes reactive and proactive nudges to ease people in, and Poletto's openly skeptical of voice-driven transactions. The goal is to take the personalized advice that today only the 20% who visit branches get, and bring it to the 80% who only ever touch the app.

18. Lenar Kess 00:08:24

    Two more quick ones from the OpenAI side. Stephanie, a solutions engineer, demoed ChatGPT inside Excel with what they call Skills — reusable templates that bake in an institution's formatting and calculation rules so analysts stop re-typing the same instructions. The pitch was traceability: the generated workbook has auditable formulas and source comments, not opaque outputs. And Connor Spicer said their coding agent, Codex, is at four million weekly active users, and that OpenAI's own engineers ship 50% more pull requests with it. Take those as vendor numbers, unaudited.

19. Damra Vol 00:09:01

    Now hold those against a contrast I want to bring in. Same day, a guy posts on the Claude subreddit — he works at Depureco, an Italian industrial vacuum manufacturer. They built a remote Model Context Protocol server over HTTP for their product catalog and wired it to Claude. No billion-dollar platform, just one company's catalog. And his takeaway wasn't that the answers got longer. It was that the system got more humble. Ask it about flour dust in a bakery, in a Zone 22 explosive-atmosphere environment, with continuous use — and instead of confidently barking 'get an explosion-proof vacuum,' it asks clarifying questions and points to the certified solution. For titanium powder off a 3D printer, it distinguishes the conductive dust class and recommends an inert rig.

20. Lenar Kess 00:09:51

    So the small deployment and the big conference make the same argument from opposite ends. The value isn't the model getting smarter — it's the model getting wired to the specific catalog, the specific compliance rule, the specific calculation standard. Grounding, plus a deterministic wrapper. The bank calls it Skills and squadlets; the vacuum engineer calls it an MCP server. Same move.

21. Damra Vol 00:10:14

    And the same risk sits underneath it, which is the handoff to the next story. Every one of these setups is a config file or a connector that tells an agent what to run and where to look. Which is exactly the surface somebody just turned into a worm.

22. Lenar Kess 00:10:28

    This is the one that should make every developer listening check their settings files tonight. There's a writeup going around — it surfaced on the programming subreddit, the analysis is from a firm called safedep — on something they're calling the Miasma worm. The headline claim is unsettling: opening a cloned repository — not installing it, not running it, just <emphasis>opening</emphasis> it — is no longer safe.

23. Damra Vol 00:10:52

    Let me walk the mechanism, because it's clever in a way that matters. The classic supply-chain attack needs a malicious dependency — you have to install something poisoned. Miasma doesn't. It hides one dropper, a file at dot-github slash setup-dot-js, and then wires that single dropper into seven different config files that tools execute on their own.

24. Lenar Kess 00:11:15

    Name them, because the list is the whole point.

25. Damra Vol 00:11:18

    Claude Code: a dot-claude settings file with a session-start hook that runs the dropper the moment an agent session opens. Gemini's command-line tool: the same structure in its settings file. Cursor: a rules file marked always-apply, so the execution instruction gets injected straight into the agent's context. VS Code: a tasks file set to run on folder-open. Then the package managers — npm's package file with the test script swapped out, Composer's post-install hook, and for Ruby's Bundler, a line of system-call code sitting on line one of the Gemfile. One commit, seven launchers.

26. Lenar Kess 00:11:58

And the trigger surface is what breaks my mental model. Some of those need a trust prompt — VS Code folder-open after you accept workspace trust, an agent session after you accept folder trust. But others have no trust gate at all. A dependency install fires the package-manager hooks. And the agent config files mean that opening the repo in your assistant is itself an execution event.

27. Damra Vol 00:12:20

    And they engineered it to hide. The dropper is a 4.3-megabyte obfuscated file. That size isn't laziness — GitHub stops indexing files for code search above roughly 384 kilobytes, so going big keeps it out of search results. The original commit, in a package called mantine-datatable, was unsigned and authored under a spoofed GitHub Actions account, with an innocent-sounding title. And the malicious bits were spread thin across all those harmless-looking config files, so no single diff looked alarming.

28. Lenar Kess 00:12:53

    And what does it actually take once it runs?

29. Damra Vol 00:12:56

    Once the loader decrypts its second stage, it scans for cloud and developer secrets — AWS, Azure, Google Cloud, Vault, Kubernetes, npm, and GitHub credentials — and exfiltrates them to public GitHub repositories the attacker creates. Which is also how it spreads: stolen GitHub tokens let it plant the same dropper in the next victim's repos. [tsk] It's a worm because the loot is the propagation vector.

30. Damra Vol 00:13:24

    And tie it back to the banks. Everything in that last segment — the Skills files, the agent connectors, the squadlet shipping straight to prod — assumes the config that steers your agent is benign. Miasma is the proof that your dot-claude folder and your dot-cursor rules are now executable attack surface, the same as a shell script. We spent a decade teaching people not to pipe curl into bash. The agent config file is the new version of that, except it fires the moment you open the repo.

31. Lenar Kess 00:13:53

    The practical takeaway is narrow, and I'll state it plainly: cloning a repo to read it isn't a read-only act anymore if you open it in an agent or an editor that honors these files. Read the safedep writeup before you next clone something from a stranger. I don't have independent confirmation of how widely this actually spread in the wild — I'm taking their account of the mechanism, which is specific and checkable, and leaving the spread claims as their reporting.

32. Lenar Kess 00:14:19

Let's shift to who owns the upside. Bernie Sanders wrote in the New York Times last week, asking more or less whether the future of humanity should be decided by a handful of AI billionaires with no democratic input. His proposed answer: a US sovereign wealth fund, built by taking 50% stock in companies like Anthropic, OpenAI, and xAI. Today in the Guardian, Nathan E. Sanders — no relation, a data scientist at Harvard's Berkman Klein Center — and the security technologist Bruce Schneier wrote a response. Their headline: Bernie's plan is good, but we think this is better.

33. Damra Vol 00:14:54

    So they're not the usual opposition. They agree with him on the diagnosis completely — they wrote a book, Rewiring Democracy, and they say the most urgent risk from AI is the concentration of power and wealth among, in their words, tech oligarchs. They agree the public should have influence and should share the upside. They just think government equity gets you the opposite of what you want.

34. Lenar Kess 00:15:17

    Spell out the mechanism, because it's a sharp argument.

35. Damra Vol 00:15:20

    Their point: if the government owns half of Nvidia, the government now has a financial interest in Nvidia's stock going up. So when chip-export decisions come up, the fund manager inside the state has the same incentive as a private shareholder — push the sales. Public ownership entangles the state with corporate profit, and they argue it makes corporate capture of government more likely, not less. Their evidence is Norway: the world's largest sovereign wealth fund holds big oil stakes, and that ownership hasn't steered those companies toward climate action — if anything Norway's dependence on the revenue has held it back. Same critique they level at US public pension funds, where the duty to grow the money overwhelms any public-interest intent.

36. Lenar Kess 00:16:08

    So what's their alternative? They split it into two goals. For sharing the rewards, use the tool we already have — taxation. They point to Elizabeth Warren's proposed excise tax on datacenter energy use, and to the idea of an AI token tax, which has a similar effect.

37. Damra Vol 00:16:25

    And notice this connects straight back to where we opened — Warren's tax falls on datacenter energy, which is the same energy driving the water draw in those drought counties. The policy fight and the resource fight are converging on one meter.

38. Lenar Kess 00:16:39

The second goal — actually shaping how AI gets built — is where they propose what they call an AI Public Option. They're not nationalizing the labs or seizing them. Government builds and runs its own models, under democratic control, as a competitive baseline that private offerings have to match or beat. Like a public option in healthcare.

39. Damra Vol 00:17:00

    And they've got a real example, which keeps it off the whiteboard. Switzerland. There's a large language model called Apertus, built by Swiss public servants and university researchers, on appropriately licensed training data, on existing public supercomputing infrastructure powered by renewables. They're candid that Apertus doesn't beat the latest OpenAI or Anthropic models on benchmarks. But they argue it wins on transparency, sustainability, and copyright compliance — and that a credible public baseline pressures the private actors to behave.

40. Lenar Kess 00:17:34

    And they draw one distinction I want to hold onto, because it sets up the next story. They say: don't confuse public AI with sovereign AI. Public AI is government building under democratic control. Sovereign AI — the idea that every country must build its own domestic AI infrastructure — they call a marketing scheme for big tech: it demands public investment without guaranteeing public control.

41. Damra Vol 00:17:58

    There's also a reporting piece from Axios today, Dan Primack, on Trump pursuing US government stakes in AI companies with, as they put it, a dealmaker's eye rather than a populist one. I couldn't get the full text behind their wall, so I'm working from the headline and summary, not the body. But the alignment Schneier and Sanders flag is exactly that — Bernie and Trump arriving at the same equity idea from opposite directions. And the authors' warning is to ask why the AI billionaires are fine with it. Their answer: because for every dollar of stock they cede, they expect more back in favorable policy.

42. Lenar Kess 00:18:37

    It's the rare policy piece that argues against a tool by reasoning about incentives instead of ideology. Whether you buy the Public Option or not, the Norway comparison is what I'll carry — owning the asset can buy you the asset's interests, not the other way around.

43. Lenar Kess 00:18:53

    And right on cue, sovereign AI showed up in the wild today — twice. South Korea's Ministry of Science and ICT put out two notices. One: they selected the operators for a government chip program worth about 2.08 trillion won — roughly a billion and a half US dollars — to procure, build,

and run GPU capacity. Two: the vice prime minister, Bae Gyeong-hoon, met with Jensen Huang of Nvidia.

44. Damra Vol 00:19:20

    And separately, Nvidia's own blog ran a post by Anthony Hills titled, more or less, how the UK is turning sovereign AI ambition into action with Nvidia technologies — a one-year-anniversary piece on the declaration Jensen Huang and Keir Starmer made at London Tech Week. So in a single day you've got a state buying chips at national scale, a deputy head of government taking a meeting with the chip vendor, and the chip vendor publishing the sovereign-AI gospel for a second country.

45. Lenar Kess 00:19:49

    And this is where the Schneier and Sanders distinction earns its keep. Strip the word sovereign off and ask the plain question: who controls the capacity after the public pays for it? A government chip fund can be genuine public infrastructure, or it can be public money buying private hardware that private firms operate and price. The notice doesn't tell us which.

46. Damra Vol 00:20:10

    Right, and I don't want to be cynical about the Korean program by default. A state that wants domestic compute for its researchers and startups has a legitimate reason to pool buying power — GPUs are scarce and expensive, and one national procurement can get far better terms than a hundred labs bidding separately. That's a genuine argument for it. The question the op-ed sharpens is narrower: when Nvidia is on one side of every one of these deals, whose sovereignty are we actually buying?

47. Lenar Kess 00:20:41

    I'll be straight about the limit of my read here: these are two press notices in Korean and one vendor blog post. I don't have the procurement terms, and I can't tell whether the capacity ends up publicly operated or leased back from a vendor. So I'm flagging the pattern, not grading the policy.

48. Damra Vol 00:20:58

    And the pattern is the one that ran through the construct show on Friday — capacity as the product. A financed asset, a delivery promise, and now a line item in a national budget. What changed today is that the buyer is a government, and the word on the box is sovereignty. The detail that would tell us which kind of program Korea's actually running is the operating model in the contract — whether the state runs the clusters or rents them back. That's the line between public infrastructure and a subsidy with a flag on it.

49. Lenar Kess 00:21:29

Two papers to close on, both on arXiv today, both about a kind of honesty in how we evaluate these systems. The first is called Attack Selection in Agentic AI Control Evaluations, and the finding is uncomfortable for anyone leaning on safety evals.

50. Damra Vol 00:21:45

    Set up the framework first, because control is a specific thing. AI control is the safety setup where you deploy a capable but untrusted agent under the watch of a weaker, trusted monitor, plus a limited human audit budget — you can only afford to hand-inspect, say, 1% of actions. You stress-test it by having a red-team attack policy try to slip something past the blue-team monitor.

51. Lenar Kess 00:22:10

    And the paper's move is to give the attacker one capability the standard evals leave out: choosing when to attack. They split it into a start policy — decide when to strike — and a stop policy — decide when to abort an attack in progress. Then they measure across two agentic environments, BashArena and LinuxArena.

52. Damra Vol 00:22:30

    And the numbers aren't subtle. At a 1% audit budget, just adding the start policy — strategic timing — drops measured safety by 20 percentage points on both environments. The stop policy drops it 20 points on BashArena and 28 on LinuxArena. Here's what matters: the attacker's underlying capability didn't change at all. Same attacks. They just picked their moments. The authors flag these as upper bounds, but the implication is direct — control evaluations that assume an attacker who strikes indiscriminately hand you an optimistic number.

53. Lenar Kess 00:23:04

    Their recommendation is concrete: future evaluations, system cards, and safety cases should elicit attack selection — make the red team choose its timing — instead of quoting a safety figure as if it held against a patient adversary.

54. Damra Vol 00:23:17

    It pairs almost too neatly with the second paper, a position paper — Stella Biderman, Naomi Saphra, and co-authors — titled Don't Just Fix It in Post.

55. Lenar Kess 00:23:27

    Read the opening line, because it's good.

56. Damra Vol 00:23:30

They open with this: 'Models are not static objects: they are snapshots of time-evolving processes shaped by data, objectives, architectures, and optimization dynamics.' Their argument is that most AI research studies the finished model — pokes at the snapshot — when the science we actually need is about how behaviors emerge during training. Fixing it in post, after deployment, is the wrong default. They lay out three rungs: prediction, forecasting an outcome from early training signals; intervention, correcting a bad trajectory mid-training; and design, building training procedures that reliably produce the behavior you want. Scaling laws already let us predict loss. Extending that to capabilities, bias, robustness, and safety is the open problem.

57. Lenar Kess 00:24:20

    And the two papers are making the same argument at different layers. The control paper says your evaluation of the finished agent is optimistic because you modeled a dumber adversary than you'll face. The position paper says evaluating the finished model at all is starting too late — the explanation lives in the training run you never instrumented. Both are arguments against trusting the static snapshot.

58. Damra Vol 00:24:43

    Which is more or less where the Friday episode ended up — the static snapshot lies. What a system is at token zero doesn't tell you what it becomes three steps in. These two papers are putting method under that intuition: one for adversaries, one for training.

59. Lenar Kess 00:24:59

    I won't force a grand unifying theory on five separate stories, so I'll just name the one element that actually recurs. In every one of these — the drought siting, the bank rewriting its platform, the worm in the config file, the public-option debate, the two eval papers — the consequential part has moved out of the model and into the system wrapped around it. And our tools for trusting that system are running behind the system itself. That's the gap I'll be reading for tomorrow. I'm Lenar Kess.

Hosts on this episode

- Lenar Kess moderator
- Damra Vol critic