When a Week Away Looked Like a Bad Score

2026-07-15 / 00:25:10

> *“When a system treats missing work data as poor work, protected leave can become a penalty without anyone writing that rule down.”*
>
> — from this episode's transcript

- Lenar Kess
- Damra Vol

A lawsuit against Meta asks whether missing work data during protected leave became a negative score, and who is accountable when people still approve the final list.

- The Guardian’s report on the Meta lawsuit lays out the workers’ allegations, Meta’s denial, and the July 22 deadline that makes the dispute urgent.
- The Guardian’s report on the publishers’ Google suit follows the alleged path from books supplied for search and sales into commercial Gemini training.
- Techmeme’s summary of Reuters reporting on xAI’s turbines connects data-center growth to permits and the communities beside the generators.
- Techmeme’s PJM power-cost item puts a $6.3 billion estimate on the capacity costs attributed to data-center demand.
- Techmeme’s summary of the data-center power pledge covers the voluntary effort to assign new generation, grid work, and unused reserved capacity to the companies creating the demand.
- Tailscale’s security bulletin explains how a username beginning with a dash could be parsed as a command option and open a root session.

- CNBC's ASML earnings report gives an upstream demand check through guidance and planned lithography capacity.
- The PalmClaw paper replaces long screen-control sequences with explicit phone tools and reports the resulting task and latency measurements.
- RunAnywhere's benchmark-runner announcement adds the useful ability to test a chosen model on the phone that will run it.

SEGMENTS

00:00:04	The absent week in the score
00:04:38	A deadline before the evidence
00:07:46	Books supplied for one purpose
00:11:42	Who receives the power bill
00:16:12	A dash that became root
00:18:21	ASML adds room for more tools
00:20:34	The agent loop moves onto the phone

Transcript

1. Lenar Kess 00:00:04

   Twenty-six current and former Meta employees sued the company this week, alleging that internal artificial-intelligence systems helped score and select workers for layoffs. The plaintiffs include people who took maternity or medical leave and people who received disability accommodations, according to the Guardian. Meta disputes the allegations and says its workforce decisions were made by people. Start with one ordinary-looking input: a person is away on approved leave, so the usual productivity signals aren't there. What does a ranking system do with the blank?

   - theguardian.com
   - theguardian.com
   - techmeme.com

- techmeme.com
- techmeme.com
- tailscale.com
- cnbc.com
- arxiv.org
- x.com

2. Damra Vol 00:00:39

   It can preserve the blank as missing data, adjust for the leave, or convert absence into a lower score. The complaint alleges the third outcome, which is why this is so uncomfortable. Nobody has to add a field that says maternity leave equals minus ten. A system can reach a similar result by expecting a stream of activity and treating a gap as evidence that the person contributed less. The protected leave remains legally approved, while the measurement layer behaves as though the employee simply stopped producing.

3. Lenar Kess 00:01:13

   According to the Guardian, the complaint points to performance ratings along with keystroke and activity-monitoring data. It alleges that Meta used those systems to, in the complaint's words, "score, rank and select employees" for a reduction of about eight thousand jobs earlier this year. The workers say their metrics either disappeared or fell while they were away. One scientist was notified two days before giving birth. An engineer says an injury leave lowered his rating, and a manager says he was dismissed sixteen days into medical leave.

4. Damra Vol 00:01:47

   Those examples make the data problem legible. A dashboard can show a neat comparison between workers while hiding that the rows weren't produced under comparable conditions. One person had a full measurement period. Another was absent because the employer had approved protected leave. If both are ranked as though they had the same opportunity to generate signals, the system has already made a policy choice before a manager sees the list. The choice is embedded in the denominator, the missing-value rule, or whatever normalization happened upstream.

5. Lenar Kess 00:02:23

   And the allegation reaches beyond a single score. The complaint describes a collection of internal systems. The Guardian separately reports that Meta introduced employee monitoring earlier this year, capturing mouse activity and browser history along with messages and location data on company devices. More than sixteen hundred employees signed a petition opposing that program, and Mark Zuckerberg paused it in June. The lawsuit asks the court for an independent audit of the

tools and the selection process. That audit would have to reconstruct which data existed, which data was absent, and how each stage used it.

6. Damra Vol 00:03:00

   An audit also has to distinguish observation from judgment. Recording a keystroke count is one act. Turning the count into a performance signal is another. Ranking people with that signal is a third, and approving terminations is another decision again. If those steps get collapsed into “the AI picked them,” you lose the exact place where a leave adjustment could have been made. If they get collapsed into “a person signed off,” you lose the possibility that every person downstream received a list whose assumptions were already fixed.

7. Lenar Kess 00:03:33

   That distinction is the center of the case. Meta told the Guardian, “Workforce management and organizational decisions were and are made by people, not AI.” I believe the second clause can be literally true and still leave the plaintiffs’ allegation unresolved. A human can make the formal decision from a ranking produced by software. The legal and factual dispute is about how much the ranking constrained that judgment, whether its inputs penalized protected leave, and whether managers could see or correct the problem.

8. Damra Vol 00:04:05

   Yes, and I don’t think “human in the loop” settles anything by itself. A person can be a reviewer, an editor, a rubber stamp, or the only one with authority to reject the recommendation. Those are different arrangements. [pause] The people in this case need the court to look at the arrangement Meta used, because the company’s denial speaks to final authority while the complaint speaks to the systems that assembled and ordered the candidates. Both statements can sit next to each other until the evidence shows how the list moved.

9. Lenar Kess 00:04:38

   The plaintiffs are scheduled to stop being Meta employees on July 22, according to their lawyers, and they’re asking the court to preserve their jobs while the dispute proceeds. That makes the timing unusually sharp. The complaint can take months or years to test, but health coverage, leave rights, unvested equity, and immigration status can change next week. How should a court handle a disputed automated process when the employment consequence arrives before anyone has inspected the process?

10. Damra Vol 00:05:08

    The request for an audit makes more sense in that light. The plaintiffs aren’t only asking for damages after the fact. They’re saying the scoring machinery contains evidence that could disappear

behind trade-secret claims, internal access controls, or routine system changes while their terminations become final. Preserving employment and preserving records are connected. If the disputed mechanism can't be examined until everyone is gone, then the company's normal timetable decides part of the case before the judge does.

11. Lenar Kess 00:05:39

    There's also a theory-of-mind problem for managers. Imagine that you're told a reduction target has to be met and you receive a ranked pool. The software presents the ordering with decimals, confidence bands, or some other appearance of precision. You may still own the decision, but the system has changed what feels like a defensible departure from the recommendation. A manager who promotes someone near the bottom now needs a reason. A manager who accepts the order may need only to say that the process was consistent.

12. Damra Vol 00:06:09

    Consistency can be the trap. Applying the same calculation to everybody sounds fair until you notice that protected leave changes the meaning of the inputs. The engineer with three missing weeks isn't a lower-output version of the engineer with a complete quarter. They're a person whose comparison period requires an adjustment. [tsk] Software is excellent at repeating a rule. It has no independent way to know that repetition has become discrimination unless the people who define the metric encode that distinction and then test whether it survives the full pipeline.

13. Lenar Kess 00:06:42

    The lawsuit hasn't established that Meta's systems worked this way, and Meta says the claims lack merit. The next evidence should show how the scoring documentation handled missing periods and what ranked outputs managers saw. It should also show which recommendations managers changed and the statistical result for workers on protected leave. A ruling on the emergency request may arrive before the broader case answers all of that. The July 22 date means the first decision will concern preservation and immediate harm, not a final verdict on algorithmic management.

14. Damra Vol 00:07:16

    And that first decision matters because it determines whether these twenty-six people can contest the system while they still have the benefits and status attached to their jobs. The complaint has taken an abstract question about automated employment decisions and attached it to a week on the calendar. By next Wednesday, either the court will have interrupted Meta's timetable or the alleged scoring process will have produced consequences that an eventual damages award can't fully reverse.

15. Lenar Kess 00:07:46

Hachette, Cengage Learning, Elsevier, and the author Scott Turow filed a separate suit against Google. They allege that books supplied for Google Books, Play Books, and Scholar were copied to train Gemini without permission. The Guardian's report gives this case a concrete chain of custody. Publishers handed Google works for defined services such as searchable snippets and ebook sales. The complaint says Google then repurposed those copies for a commercial model.

16. Damra Vol 00:08:14

    That history changes the argument. The parties aren't starting with a mystery archive whose origin has to be inferred. They had a relationship, the books entered Google's systems through named products, and those products came with permissions the publishers say were limited. The court may have to ask whether indexing, displaying a snippet, selling a digital copy, and training a model are uses covered by the same grant. Permission to possess a file isn't automatically permission for every computation you can perform on it.

17. Lenar Kess 00:08:45

    The complaint names works including N. K. Jemisin's The Fifth Season and Lemony Snicket's Who Could That Be at This Hour? It seeks statutory damages, an injunction, and destruction of unauthorized training copies. Google hadn't responded to the Guardian's request for comment when the report was published. The publishers also allege that internal discussions recognized large legal exposure, though those claims haven't been tested in court.

18. Damra Vol 00:09:12

    Destruction is the technically strange remedy here. If a court accepted the allegation, what exactly counts as the unauthorized copy once training has happened? The source files may exist in a dataset, a preprocessed corpus, backups, and intermediate artifacts. Then there's the model itself, which isn't a library of readable books but may still carry information derived from them. A deletion order sounds simple in a complaint and becomes a systems question as soon as someone has to prove compliance.

19. Lenar Kess 00:09:44

    The Guardian quotes the suit saying a generated hundred-page murder mystery could cost thirty-nine cents and compete with books used in training. That's the publishers' substitution argument, and it's an aggressive one. A cheap generated story isn't necessarily a substitute for a named author's novel. Still, the allegation identifies the economic fear cleanly: Google may have received books for distribution services, used them to improve a product, and then offered output that competes for some of the same reading time.

20. Damra Vol 00:10:14

I’m more persuaded by the permission boundary than the twenty-minute murder mystery. The generated-book example invites a fight over quality and consumer taste. The prior agreement asks a narrower question: what use did the publishers authorize when they supplied the files? If the documents and contracts support their account, Google can’t answer only with a general theory of transformative machine learning. It also has to explain why material acquired for one service could be moved into another commercial purpose.

21. Lenar Kess 00:10:47

    That keeps today’s case distinct from yesterday’s Australian policy argument. This filing is about Google’s own possession of particular books and the permissions attached to named services. The pleadings will matter more than slogans about authors versus models. Contracts, product terms, internal transfer records, and the technical lineage of the training set can tell the court whether the alleged repurposing happened and what Google believed it was allowed to do.

22. Damra Vol 00:11:13

    It also means a future response from Google has several available routes. Google could dispute that the books entered Gemini training, argue that the existing permissions covered the use, or defend the training as lawful even without a license. Each answer produces a different case. The complaint has supplied the publishers’ route through the building. Google’s account of that route will determine whether this becomes a contract-heavy dispute, a fair-use dispute, or both.

23. Lenar Kess 00:11:42

    Three reports this week put numbers and equipment around data-center electricity costs. Reuters reporting summarized by Techmeme says xAI installed more gas turbines at its Colossus 2 project than it had publicly acknowledged and lacked federal clean-air permits for them. A separate report attributes about six point three billion dollars in a PJM capacity auction to data-center demand. Another describes a voluntary pledge meant to keep those costs off existing utility customers.

24. Damra Vol 00:12:12

    Those are three different exits from the facility. The turbines move generation beside the data center, so the questions become permits, emissions, and who lives nearby. The PJM auction spreads capacity costs across a regional market that serves thirteen states and Washington, D.C. The pledge tries to move generation and grid-upgrade costs back toward the companies creating the new demand. You can support rapid construction and still argue fiercely about which door the bill is allowed to leave through.

25. Lenar Kess 00:12:45

The xAI reporting needs attribution because the underlying account is disputed territory and today's item reaches us through a summary. The claimed mechanism is straightforward, though. On-site gas turbines can supply power before a full grid connection is ready. That can shorten the wait for compute, but permits don't disappear because the electricity is generated behind the facility fence. Reuters' analysis, as summarized by Techmeme, says the emissions fall hardest on nearby Black communities.

26. Damra Vol 00:13:15

    And the fence doesn't contain the air. That's why "bring your own power" is incomplete as a social answer even when it works as an engineering answer. A facility can avoid drawing every megawatt from the grid and still impose pollution, noise, water use, or transmission work on its neighbors. The local permit is where a claim about private infrastructure meets people who didn't sign the compute contract.

27. Lenar Kess 00:13:41

    PJM exposes a different accounting problem. The report puts the capacity auction at sixteen point four billion dollars. The grid monitor attributes roughly six point three billion of that to data centers. The auction also fell short of the future supply PJM wanted to secure. So customers face high capacity payments while the region still has less reserve than the target. The proposed answer is a dedicated process that would assign more of the added generation cost to data centers.

28. Damra Vol 00:14:11

    Ratepayers can face a higher bill without the comfort of knowing the system bought enough spare capacity. They don't experience an auction design. They experience a monthly charge and a summer warning about demand. If data centers need an unusually large block of new supply on an unusually fast schedule, separate procurement can make the causal link visible. It can also reveal whether the promised demand is firm enough for developers to finance generation against it.

29. Lenar Kess 00:14:40

    The White House effort described in the reporting is voluntary. Large technology companies have already pledged to finance new generation, grid work, and unused reserved capacity tied to their projects. The next event is expected to bring in utilities, data-center developers, and governors. The open issue is implementation. A promise that current customers won't pay needs tariffs, contracts, and commission decisions that assign each cost to somebody.

30. Damra Vol 00:15:08

    Voluntary commitments can still change negotiations if utilities begin treating them as a baseline. A developer seeking a connection may find that the governor, the commission, and the utility all ask

for a dedicated tariff before approving the project. But a pledge can also become ceremony if the expensive categories remain undefined. Who pays when a transmission upgrade serves the data center today and a whole county ten years from now? Cost allocation becomes a fight over the useful life of the asset, not only who requested it first.

31. Lenar Kess 00:15:42

    New York's pause yesterday established the policy conflict around large projects. Today's addition is the invoice. One facility can install turbines, a regional market can spread capacity payments, or a developer can contract to fund new supply. Each method gets compute online under a different bargain with neighbors and customers. The next concrete events are PJM's filing on its emergency procurement process and the permit record around Colossus 2.

32. Lenar Kess 00:16:12

    Tailscale fixed an SSH bug in version one point ninety-eight point nine that could turn a permitted non-root connection into a root session. Its bulletin says the server accepted usernames beginning with a dash and passed them to the Linux account lookup command. A username made from a dash and the letter I was interpreted as an option, causing the command to print the password database beginning with the root entry. Tailscale then opened the session as that first returned user.

33. Damra Vol 00:16:40

    [tsk] The access-control policy can be correct, and the caller can already have legitimate SSH permission. One unmarked string boundary still changes the requested identity. The username travels from network input into a command-line argument. Because the receiving program treats a leading dash as control syntax, data becomes an option. Then a second assumption takes the first returned account as the answer. Two familiar behaviors compose into root access.

34. Lenar Kess 00:17:14

    The scope matters. Tailscale says this affected Linux hosts using its SSH feature and relying on the non-root group restriction. It wasn't an unauthenticated internet-wide root bug; the attacker already needed Tailscale SSH access to the node. Within that boundary, though, the policy violation was severe. The fix rejects usernames with leading dashes, and the vendor says users of Tailscale SSH should upgrade to one point ninety-eight point nine or newer.

35. Damra Vol 00:17:44

    I like the bulletin because it names the whole sequence without fog. A username enters as input, the parser reads it as an option, and Tailscale accepts the first line of output as the account. You can reproduce the bug in your head. It's also a reminder that identity code touches old Unix conventions in surprising places. The dangerous step wasn't an exotic cryptographic break. An identifier crossed

into an argument position without ending option parsing or rejecting impossible account names. Tailscale chose rejection, which is easy to explain and test.

36. Lenar Kess 00:18:21

    ASML raised its annual sales guidance for the second time this year after stronger-than-expected quarterly results, according to CNBC. The company said customers are accelerating production expansion. It plans to add about thirty percent to its 2026 capacity for low numerical-aperture extreme-ultraviolet lithography systems. Deep-ultraviolet immersion capacity will also rise by thirty percent. Those machines sit several steps upstream from the AI systems people touch.

37. Damra Vol 00:18:52

    That makes the order book useful and limited. ASML is the only supplier of the extreme-ultraviolet machines required for leading-edge chips, so chipmakers reserving more equipment tells you they're preparing to manufacture more advanced logic and memory. It doesn't tell you that every future model service will find paying users. Equipment orders are a commitment to capacity, made years before the final demand is visible. They're evidence that the manufacturers still believe the buildout will need a lot more wafers.

38. Lenar Kess 00:19:24

    The deep-ultraviolet increase is useful too because the AI supply chain isn't made only of the most advanced accelerator die. The supply chain also depends on memory and networking. Power-management chips, controllers, and other supporting components use different processes. More immersion-tool capacity suggests customers expect pressure across a broader set of components. Export controls remain a constraint on which systems ASML can sell into China, but the announcement today is mainly a production decision: the company is preparing its own factories and suppliers for higher equipment output.

39. Damra Vol 00:20:01

    And that is less reversible than an enthusiastic earnings-call adjective. Adding capacity for machines this complex means training people, qualifying suppliers, and creating room for assemblies that operate at absurd precision. [chuckle] You don't do that because one chatbot had a good quarter. You do it because several chipmakers are showing road maps and purchase plans that occupy the same years. The uncertainty hasn't vanished; ASML is choosing to carry more manufacturing capability into it.

40. Lenar Kess 00:20:34

    A new paper called PalmClaw puts the agent framework on an Android phone and exposes phone capabilities as explicit tools instead of controlling everything through taps, swipes, and screenshots.

The authors run sessions, memory, skills, tool execution, and the agent loop on the device. The language model is still reached through a remote API, so this isn't fully local inference. The orchestration and the phone actions are local.

41. Damra Vol 00:21:01

    That division is interesting because the phone stops being a puppet screen attached to a desktop agent. A calendar action can be a structured request with a date, time, and permission check. A file action can resolve against an approved workspace. The model may live elsewhere, but the device owns the execution boundary. That gives the operating system a place to ask for confirmation and gives the user a record of which capability the agent invoked.

42. Lenar Kess 00:21:29

    The paper evaluates seventy mobile tasks and a nineteen-task subset of AssistantBench. On its MobileTask set, PalmClaw reports a ninety-seven point one percent success rate. The strongest baseline reached eighty-seven point one percent. Average completion time falls from three hundred forty-eight point eight seconds to seventeen point seven seconds. The authors summarize the results as an eleven point five percent relative improvement in success and a ninety-four point nine percent reduction in time.

43. Damra Vol 00:22:00

    The action count explains much of the speed. PalmClaw averages two point eight model calls on those tasks. The screen-driven systems can require dozens of actions because they must observe an interface, locate a control, tap it, and then inspect the next screen. A calendar tool can express the intended operation in one bounded call. That's a comparison between action spaces as much as a comparison between agent frameworks, and the paper uses the same DeepSeek model across systems to keep the language model from swallowing the result.

44. Lenar Kess 00:22:36

    There are still reasons to treat the benchmark as an early result. The authors constructed MobileTask from existing sets, used one Android phone, and used the same model as a judge for some answer-based tasks. Their tool approach also avoids interface work that a general screen agent can attempt in any app. PalmClaw reports unsupported operations when no tool exists. That narrower capability set is part of why the actions are faster and part of what a user would have to evaluate.

45. Damra Vol 00:23:06

    The limit makes the product choice visible. A phone agent with ten dependable tools may be more interesting than one that can theoretically tap anything and often gets lost. The paper's trace

examples show a calendar permission request and a confirmation before taking a photo. A file write also stops at the workspace boundary. Those moments tell you what the agent can't silently do. Generality is exciting, but a device full of personal data benefits from a capability list you can inspect.

46. Lenar Kess 00:23:38

    RunAnywhere announced a benchmark runner alongside this that measures a selected model on the user's own device. That pairing is practical. A framework can tell you how it organizes tools and permissions, while the runner can tell you whether a model is fast enough on the phone you intend to use. "On-device" covers a huge range of processors, memory limits, thermal behavior, and operating-system policies. A number from the developer's newest phone may say little about the handset in someone else's pocket.

47. Damra Vol 00:24:08

    Phones also change the emotional register of agents. A server agent acts on a remote workspace. A phone agent is beside your conversations, photos, location, calendar, and microphone. A phone-native agent should express its actions in the operating system's permission language, run fast enough on the actual hardware, and stop unsupported requests before improvisation turns into access.

48. Lenar Kess 00:24:33

    PalmClaw now needs evidence from more phones and models, along with independent task sets and longer use where memory and scheduled wakeups matter. RunAnywhere's runner can help make some of those hardware differences measurable. The paper has shown a credible alternative to watching an agent peck at a screen: keep the decision loop close to the device, give each action a typed interface, and let Android enforce the moment when a calendar, camera, or file crosses into personal space. Lenar Kess.

Hosts on this episode

- Lenar Kess moderator
- Damra Vol critic