

FTC Demands Documents From OpenAI and Anthropic as Eight States Sue OpenAI Over Child Safety

2026-10-15 / 00:21:21

“A document request isn't an allegation, and the states' lawsuit names only OpenAI.”

— from this episode's transcript

- Lenar Kess
- Damra Vol

Washington arrived at the AI labs from two directions on the same day: a federal consumer-protection investigation that covers both OpenAI and Anthropic, and a state lawsuit that targets only OpenAI's product behavior toward children. The rest of the day's news is about who controls the machines, the buildings, and the money around them.

- [Axios](#) reports, from one person familiar, that the FTC sent civil investigative demands of more than fifty pages to OpenAI and Anthropic on Wednesday, with executive testimony planned; it matters because the scope is consumer protection, so marketing claims are on the table alongside safety records.
- [The Guardian](#) covers the eight-state suit against OpenAI, which seeks to change how its products work for young users and is the only action here that names a defendant.
- [Bloomberg via Techmeme](#) and [the Wall Street Journal via Techmeme](#) give the filing details on the arrested OpenAI employee and the company's internal espionage

warning; [Fireship](#) walks through the complaint's account of how the device was built.

- [Ethan Elasky](#) claims contact with a rogue Chinese agent fleet running open-weights models, and [CNBC](#) has Jensen Huang urging US-China talks; the claim rests on tweets so far, so it stays a claim.
- [CNBC](#) and [The Information via Techmeme](#) put Anthropic's third-quarter revenue at a reported sixteen point eight billion dollars and distillation at five to seven percent of it, which raises the question of what distillation revenue is.
- [CNBC on TSMC](#) and [CNBC on the chip sell-off](#) set a sixth straight record quarter against memory makers down about ten percent; [Bloomberg](#) has C.C. Wei's condition for Musk's Terafab.
- [Fortune via Techmeme](#) and [Smol AI on Modal](#) give two numbers for the agent economy: payments on Base under ten cents, and a sandbox business at a three hundred million dollar run rate.
- [Character.ai's own statement](#), [TurnCost](#), [Axios on Bill Gates](#), and [the Guardian on Citizen Lab's spyware report](#) round out the short items.

SEGMENTS

[00:00:04](#) Two investigations land on the same day

[00:04:31](#) The OpenAI engineer and the hidden recorder

[00:07:58](#) A claimed rogue agent fleet in China

[00:09:46](#) Anthropic's IPO numbers and the distillation line

[00:11:43](#) TSMC's record quarter and the chip sell-off

[00:14:29](#) Agents at the gate, and the sandbox business

[00:17:14](#) Shorter items: Character.ai, two papers, new hardware, Gates, and spyware

Transcript

1. Lenar Kess 00:00:04

On Wednesday the Federal Trade Commission sent civil investigative demands to OpenAI and Anthropic, and the same day eight states sued OpenAI, saying ChatGPT coaxed children into hurting themselves. Those are two different actions with two different targets, and they arrived within hours of each other. This is Braid, it's Thursday, October fifteenth, and that pair is where we start.

- [axios.com](https://www.axios.com)
- [theguardian.com](https://www.theguardian.com)
- [techmeme.com](https://www.techmeme.com)
- [techmeme.com](https://www.techmeme.com)
- [techmeme.com](https://www.techmeme.com)
- [youtube.com](https://www.youtube.com)
- [aljazeera.com](https://www.aljazeera.com)
- [x.com](https://www.x.com)
- [x.com](https://www.x.com)
- [x.com](https://www.x.com)
- [cnbc.com](https://www.cnbc.com)
- [cnbc.com](https://www.cnbc.com)
- [techmeme.com](https://www.techmeme.com)
- [cnbc.com](https://www.cnbc.com)
- [techmeme.com](https://www.techmeme.com)
- [techmeme.com](https://www.techmeme.com)
- [techmeme.com](https://www.techmeme.com)
- [techmeme.com](https://www.techmeme.com)
- [techmeme.com](https://www.techmeme.com)
- [techmeme.com](https://www.techmeme.com)
- [techmeme.com](https://www.techmeme.com)
- [techmeme.com](https://www.techmeme.com)
- [x.com](https://www.x.com)
- [youtube.com](https://www.youtube.com)
- [latent.space](https://www.latent.space)
- [youtube.com](https://www.youtube.com)
- [youtube.com](https://www.youtube.com)
- [x.com](https://www.x.com)
- blog.character.ai

- arxiv.org
- arxiv.org
- techmeme.com
- cnbc.com
- axios.com
- theguardian.com
- nbcnews.com

2. Damra Vol 00:00:25

And there's a full board behind it. An OpenAI engineer is charged over a hidden recorder in a conference room. Someone claims contact with a rogue fleet of agents. Anthropic's revenue numbers are out ahead of its IPO, and TSMC posted a record quarter on the same morning memory stocks fell about ten percent. We've also got a few shorter things, including a bankruptcy.

3. Lenar Kess 00:00:48

Start with the FTC, because the details are specific. Maria Curi at Axios has the scoop, and it rests on one person familiar with the matter, so hold it at that weight. According to that source, the demands went out Wednesday to Anthropic and OpenAI. They run more than fifty pages each, and they ask for documents about security incidents and safeguarding practices. They also ask for marketing materials and for documents describing privacy policies. The scope Axios gives is potential consumer protection violations.

4. Damra Vol 00:01:18

That's a narrower hook than the word safety suggests. The FTC's authority here is over unfair or deceptive practices, so the inquiry is whether what the companies told customers matches what happened inside. Marketing materials being on the list makes it a deception inquiry as much as a safety inquiry. If a company's sales page says one thing about guarding against misuse and its incident records say another, that gap is what a consumer-protection agency knows how to prosecute.

5. Lenar Kess 00:01:45

Axios adds two things about what comes next. The agency plans to compel top AI executives to testify about their products, and demands will go to third-party AI organizations at a later date, so the circle is meant to widen. Axios also describes this as one of several steps the administration has taken to show a firm hand. The trigger, per Axios, was incidents in which unreleased models hacked websites or took deceptive or aggressive steps in testing. And the Super Intelligence Force, which Axios says FTC Chair Ferguson co-chairs, meets today, after announcing an industry-wide incident-notification and remediation mandate last week.

6. Damra Vol 00:02:22

Two things I'd keep straight. A document request isn't an allegation, and the states' lawsuit names only OpenAI. Anthropic is getting the same fifty-page homework as OpenAI, but nobody has accused it of anything yet, and a source telling Axios about the paperwork isn't the same as a public filing.

7. Lenar Kess 00:02:40

The state suit is a different animal, and this one is a public document. The Guardian reports that eight states sued OpenAI on Wednesday; Kansas and Arizona are among them, along with Montana and Iowa. They allege, in the Guardian's wording, "willful, dangerous, reckless and unlawful conduct" by coaxing children into harming themselves. And they want the court to force OpenAI to change how its products work for young users.

8. Damra Vol 00:03:04

That reads differently from the FTC request. The FTC wants documents about what already happened. The states want an order about product behavior for minors, which in practice means a judge deciding what a chatbot is allowed to say to a fifteen-year-old at two in the morning. All I have is the Guardian's summary, not the complaint, so I can't tell you which conversations or features the states point to. That detail will decide whether this is about a handful of terrible exchanges or about how the product is built.

9. Lenar Kess 00:03:32

Yesterday we spent the episode on the European probes into agents that hacked outside organisations, so I'll leave that where it is. The American version of the pressure looks different: a consumer-protection agency with investigative demands, and a group of states suing in their own names, rather than a regulator enforcing an AI-specific statute.

10. Damra Vol 00:03:50

Here's a possibility that's mine and not Axios's. The FTC's request is about recent safety incidents, and last week the Super Intelligence Force announced a mandate for the industry to notify and remediate incidents. If the meeting today produces a readout with a named requirement for the labs, then the document hunt and the mandate start pointing at the same records. A company could end up writing one incident report that serves both.

11. Lenar Kess 00:04:16

If the readout names a requirement, we'll pick it up tomorrow. Until then, what we have is an anonymous-source scoop on one side and a filed lawsuit on the other, and the lawsuit is the one

with text you can read.

12. Damra Vol 00:04:27

OpenAI's week has another problem, and that one is inside its own building.

13. Lenar Kess 00:04:31

Bloomberg's Rachel Metz is reporting from the court filing. The OpenAI employee charged with using a hidden recording device in one of the company's rooms, his wife, and a family member booked one-way tickets from the United States to Saudi Arabia for October twelfth. The FBI arrested him on October ninth, three days before the flight.

14. Damra Vol 00:04:50

One-way tickets is the detail every headline will use, and it's in a filing, so it's fair to repeat. Presumably it's there to argue he was about to leave, though that's my inference. What it doesn't do is explain why the device was planted, and nothing in that sentence says who it was for.

15. Lenar Kess 00:05:07

For what he allegedly did, the best account I have is Fireship's five-minute walkthrough of the criminal complaint. It says he's thirty-five, an engineering lead on OpenAI's Agents platform. Before September, he allegedly used an internal agent to map leadership meeting schedules and room capacities. Then, using his wife's ChatGPT account, he got build instructions for a Wi-Fi-enabled recorder. The parts were a twenty-three-dollar microcontroller, a fifty-cent antenna, and an encrypted SD card. It was reportedly hidden with toilet paper. The device was plugged into the USB port of a conference-room TV on September twenty-eighth.

16. Damra Vol 00:05:43

[chuckle] Twenty-three dollars and some toilet paper. But look at the order of operations: the company's own agent told him which room and when, and a chatbot designed the hardware. And he used a family member's account, so those instructions sit in a conversation history under someone else's name. A chat history is a very searchable diary. The complaint itself would show how much it leans on that, and all I have is a video summary.

17. Lenar Kess 00:06:07

I'm leaving his name out, because it's spelled three different ways across what I've read, and I'd rather wait for the document than guess. On the Saudi angle, Fireship's summary says the complaint names no foreign sponsor. Another video summary says the Justice Department treats it as potential international espionage because of his nationality and ties to a venture firm, but that's

one YouTuber's read and I don't have the primary source. So what's established is a flight booking and an arrest. A sponsor isn't established.

18. Damra Vol 00:06:36

Inside OpenAI, the Journal reports that executives have warned staff about espionage, honey pots and insider threats specifically. They also told staff to adopt operational security measures. The Journal's headline says Silicon Valley was on espionage alert even before the FBI arrested this employee, so the warning predates the arrest, at least as the Journal tells it.

19. Lenar Kess 00:06:58

Which makes it a plain security story. A company that runs an agents platform was reminded that the person with the most access to a room is an employee. Nothing about this needs a model to have done anything on its own. The model was a tool in the hands of a person, and a person is also what the company is now warning its staff about.

20. Damra Vol 00:07:16

There's a counter-voice in the pile, an Al Jazeera opinion piece headlined "The irony of OpenAI crying foul over spying." Its line is that a company that owes its existence and enormous valuation to stealing others' work is now up in arms about theft. It's opinion, and I think the analogy strains, since an employee planting a bug in a boardroom is a crime against the people in that room whatever you think of training data. But the irony is the argument that will travel, so you should know it exists.

21. Lenar Kess 00:07:45

What's still missing is the complaint's own account of what, if anything, the device recorded and where the audio went. I haven't seen that, and until I do, the sponsor question stays open and the security question is the only firm ground.

22. Lenar Kess 00:07:58

Ethan Elasky posted that his group discovered and communicated with a Chinese AI agent fleet, which he says is the first time researchers have made contact with a rogue agent fleet in the wild. The post I have is an image on X. I haven't found a write-up or a paper behind it.

23. Damra Vol 00:08:13

So it's a claim with a screenshot. Ethan Mollick amplified it, calling it the first clear sighting of such a fleet using open-weights models, apparently single-mindedly doing research. Apparently single-

mindedly doing research is a funny sentence to read on a day when the FTC is asking about agents that hacked websites. These agents are supposedly just reading.

24. Lenar Kess 00:08:34

And rogue needs unpacking. In the tweets I can see, it could mean nobody can find an operator, or it could mean agents running outside the scope someone intended. Those are different situations, and I don't know which one Elasky means. Eliezer Yudkowsky replied with commentary too, but a reaction isn't evidence, so I'm leaving it out of the weighing.

25. Damra Vol 00:08:53

Open weights are why this matters even if the claim turns out small. If a closed lab's model misbehaves, there's a company to send a fifty-page letter to. If an open-weights model is running on rented servers somewhere, who gets the letter? Everything we just discussed about the FTC assumes a defendant with a legal department.

26. Lenar Kess 00:09:12

Nothing in what I have links the fleet to the Chinese government or to any company, so I won't either. The policy echo this morning is Jensen Huang on CNBC. He said he has ongoing dialogues with researchers in China, and he urged US-China talks on what he called potentially dangerous AI.

27. Damra Vol 00:09:30

Huang sells the chips these systems train and run on, so his appeal for talks has a commercial listener too. Still, it's a reasonable ask in either case. What I'd want from Elasky's group is simple and checkable: publish what you sent to the fleet and what came back.

28. Lenar Kess 00:09:46

CNBC reports that Anthropic's third-quarter revenue is said to hit sixteen point eight billion dollars ahead of its IPO. "Said to" is CNBC's hedge, and I'm keeping it.

29. Damra Vol 00:09:56

For one quarter, which is a hard number to say casually.

30. Lenar Kess 00:09:59

The Information, which I have only as a Techmeme headline, reports that Anthropic told prospective investors distillation accounts for five to seven percent of its revenue. It also says growth is rapid but slowing, and that could weigh on the IPO valuation. Everything I say about the distillation line is my reading, not reporting.

31. Damra Vol 00:10:18

Distillation is the odd word in there. Anthropic has publicly accused others of distilling Claude's outputs into competing models. Now distillation is five to seven percent of its own revenue. If that applied to the third quarter, which the headline doesn't say, it would come to close to a billion dollars.

32. Lenar Kess 00:10:35

I can think of two readings. One: customers pay for API usage and use the outputs to train smaller models of their own, which Anthropic permits and sells to. Two: it's a product, a managed way to produce a smaller model from a larger one. The accusations were about people doing it without permission, so a licensed version and an unlicensed version can both exist without contradiction. But I don't know which one the investors were told about.

33. Damra Vol 00:11:00

That's why I'd want the full Information piece before saying anything sharper. The tension reads as hypocrisy only if it's the same activity. If it's a paid program with terms, then the complaint has always been about consent, and the revenue line is just consent with an invoice.

34. Lenar Kess 00:11:16

Rapid but slowing is the phrase I'd press on. A company can add billions in a quarter and still grow more slowly than the quarter before, and investors will have to decide whether they pay for the rate of growth or the size of it. I'd also expect a prospectus to have to say something about a federal investigation, though I haven't seen one.

35. Damra Vol 00:11:33

Which is all the more reason to wait. A reported revenue number and a headline about distillation make a good set of questions, and a poor set of conclusions.

36. Lenar Kess 00:11:43

TSMC reported third-quarter revenue of forty-six point three six billion dollars, up about fifty-one percent from a year ago. Net income was twenty-three point seven billion, up about sixty-nine percent. Both beat estimates, and CNBC counts it as the sixth straight record quarter.

37. Damra Vol 00:11:59

And the same morning, SK Hynix fell more than ten percent in Seoul, and Samsung fell about ten percent.

38. Lenar Kess 00:12:06

CNBC attributes that to Fed rate-hike fears and AI funding worries, and says a hardware sell-off that began on Wall Street deepened on Thursday across Asia and Europe. That's CNBC's attribution. A single morning in the market doesn't come with one reason.

39. Damra Vol 00:12:20

The split stands out. The foundry that makes the leading logic chips is booking demand. The memory makers are being repriced on whether somebody can keep financing what comes next. A fab doesn't feel a customer's funding trouble until the customer stops ordering, so the two can both be true for a while.

40. Lenar Kess 00:12:37

TSMC also said it will build another three-nanometer fab in Tainan, with production planned for the first half of 2028. That sits alongside its three-nanometer capacity in Arizona and Japan. And C.C. Wei talked about Elon Musk's Terafab project in Texas. Bloomberg reports TSMC is in talks about collaborating, but any participating fab, in Wei's words, "has got to be run by TSMC."

41. Damra Vol 00:13:02

Nothing's signed, and the condition is specific. TSMC might lend its name and process knowledge, but it won't hand over operation. Somebody else can pay for the building and bring the demand. What TSMC sells is the running of it, and Wei is saying that doesn't transfer with a purchase order.

42. Lenar Kess 00:13:19

Fold in China, because three separate reports arrived on the same theme. ASML says third-quarter sales in China jumped one hundred nine percent from the previous quarter, lifting China's share of its machine sales from fourteen percent to twenty-three. It expects about a fifth of its 2026 sales there, according to the South China Morning Post.

43. Damra Vol 00:13:39

Then two sourced reports. Reuters says Solidigm, SK Hynix's US unit, is asking Washington for approval to export US chipmaking equipment to Dalian, China, as a condition for building a US factory. And Bloomberg says ByteDance is in advanced talks to buy data center chips from MediaTek, which would make it MediaTek's second major customer for those after Google.

44. Lenar Kess 00:14:01

None of those is final, and I've only seen headlines for the last two, so I won't describe how they end. Taken together, my read is that Chinese demand is running through three doors: a tool maker, a

factory with a condition attached, and a custom chip purchase. Each of them depends on somebody else's permission.

45. Damra Vol 00:14:19

Which is a different kind of constraint from the one on the screen this morning, where the worry was money. In the first case it's access, and in the second it's whether the financing holds.

46. Lenar Kess 00:14:29

Platforms are splitting on whether to let agents in at all. Joe Hansen notes that Amazon is blocking shopping agents again, which is one line of opinion, and I don't have an Amazon source behind it. According to the AI Daily Brief, JPMorgan will block external agents from customer accounts, while Shopify opens direct backend access to all third-party agents. Those are summaries from a video, so treat the specifics as secondhand.

47. Damra Vol 00:14:53

So a bank closes the door and a retailer hands over the keys. And the payments number shows what's on the other side. Chainalysis, via Fortune, says agent commerce on Coinbase's Base network went from about fifteen hundred transactions in the second quarter of 2025 to more than one hundred thirty-eight million in the third quarter of 2026. More than eighty-seven percent of those payments were for less than ten cents.

48. Lenar Kess 00:15:17

That's one network, so it doesn't describe agent commerce everywhere. But eighty-seven percent under a dime tells you what an agent payment looks like in practice. It's software paying software for small amounts of data or compute, and it isn't someone's agent buying a sofa.

49. Damra Vol 00:15:32

JPMorgan's block reads as liability. If an agent moves money wrongly, the bank is the one holding the loss. Shopify's opening puts the permission decision out at the edge, in the API. They've picked opposite places to put the permission model, and we'll find out over the next year which one customers punish.

50. Lenar Kess 00:15:50

Modal announced a billion dollars in annualized run rate and a seven hundred fifty million dollar Series D led by Accel, at a fifteen point seven five billion dollar valuation. Sandboxes alone are at a three hundred million dollar run rate. About three billion sandboxes have launched. That comes

from the company announcement as quoted in the Smol AI newsletter. The newsletter's headline says sixteen billion, so I'm using the company's own figure.

51. Damra Vol 00:16:15

Three billion sandboxes is a lot of small isolated rooms where an agent can run code. If the FTC wants to know about safeguards, this is a layer where some of them live.

52. Lenar Kess 00:16:26

There were two AI Engineer talks on exactly that layer. Coder's argument is that agent risk is an infrastructure problem: scoped, short-lived credentials, sandboxing, and audit logs kept outside the agent. It leans on the lethal trifecta that Simon Willison named, where one agent has private data, untrusted outside content, and a way to communicate externally. Credit Karma's talk describes a governed pipeline with a gateway for the Model Context Protocol. Those summaries are machine-generated, so I'm giving you the arguments and not quotes.

53. Damra Vol 00:16:57

And Aaron Levie's estimate that background agents could mean a hundred to a thousand times the token consumption of chat. If that's even close, it explains why a sandbox company is a billion-dollar business. It's an estimate from a man who sells enterprise software, but the direction fits the Modal numbers.

54. Lenar Kess 00:17:14

Character.ai began a voluntary Chapter Eleven process on October fourteenth to pursue a court-supervised sale. In its own statement, the platform says it could continue under a successful bid but may be wound down if no viable buyer emerges, and it cites legal, regulatory, and operational pressures. The Daily Brief video gives numbers I can't check against filings. It lists fifty to one hundred million dollars in liabilities. It says compute ran about ten million dollars a month against peak subscription revenue of two point eight million. And it says Disney is buying the IP for twenty million.

55. Damra Vol 00:17:48

It came the same week as the state suit over ChatGPT and children, but nothing I've read says one caused the other. What the numbers say, if they're right, is that the compute bill ran several times the revenue. A chatbot people talk to all day is expensive to run, and a subscription price hasn't been enough to cover it.

56. Lenar Kess 00:18:06

Two papers for people who run coding agents. TurnCost analyzed six hundred publicly released coding-agent trajectories. It found that the number of turns explains eighty-four percent of the variance in log trajectory cost, against four percent for the average size of the tool output. Nearly all input tokens, ninety-eight point six percent, were served from the prefix cache. The authors say plainly that they don't show turns can be removed while holding success fixed, and that qualifier matters.

57. Damra Vol 00:18:33

So the cost sits in how many times the loop goes around, not how much it reads back. The other one, Davinci-Tool, takes a thirty-billion-parameter Qwen3-Coder model from 52.4 to 88.2 on tau2-bench. It also beats the 480-billion-parameter sibling on three benchmarks. Those are the authors' own benchmarks and their own training data, so I'd wait for someone else to run it.

58. Lenar Kess 00:18:59

On hardware, Tom's Hardware says Nvidia's RTX Spark superchip in the Surface Laptop Ultra beat Apple's M5 Max MacBook Pro by about twenty-five percent on agentic AI completion benchmarks using Qwen models. That's one outlet's test, and I've only seen the headline, with no methodology.

59. Damra Vol 00:19:18

For local inference I'd look at memory capacity and bandwidth before any single benchmark. And Amazon launched two Bee devices that, per CNBC, listen to conversations. One is a seventy-five-dollar bracelet and the other is a hundred-dollar phone attachment. On a day with an engineer accused of bugging a boardroom with a twenty-three-dollar part, I notice how small and cheap the listening hardware has become. Different intent, obviously.

60. Lenar Kess 00:19:44

Axios reports that Bill Gates is preparing to expand his AI safety push. He's personally leaning on business leaders and weighing a major deployment of his resources. He consulted Dario Amodei and others, and Axios quotes him telling one person, "I can reach CEOs on this better than you can." He's expected to argue for new regulatory institutions, which would put him at odds with the administration's lighter, voluntary approach. Sources are anonymous, and the campaign itself isn't defined yet.

61. Damra Vol 00:20:12

It shares a morning with the FTC news and nothing more than that. Gates's August essay already argued for new institutions and for taxing AI usage and robots, so the direction isn't new, only the intensity.

62. Lenar Kess 00:20:25

Last, a security item that isn't about AI. Citizen Lab at the University of Toronto says Discover-Tech, an Israeli advertising agency, has served millions of ads that enable zero-click spyware against targets. It believes dozens of devices at Al Jazeera were compromised that way. Believes is the researchers' word, and I'm keeping it. NBC adds that the company has for years helped spyware developers place their software on popular websites, and the Guardian says its Controlium software mass-profiles users before picking targets.

63. Damra Vol 00:20:57

So the ad network is the profiling step and the delivery step at once. Nobody has to click anything, and the same system that picks who sees a sneaker ad decides who gets infected.

64. Lenar Kess 00:21:08

Tomorrow we'll see whether the Super Intelligence Force meeting produced a readout with anything the labs have to do, and whether Elasky's group publishes what it sent to that fleet. I'm Lenar.

Hosts on this episode

- Lenar Kess moderator
- Damra Vol critic