



AI GENERATED 
TRANSPARENCY DISCLOSURE

THE OTHER SIDE · A FYTAHQ COMPANION RESOURCE

Inside Anthropic's Red Lines

Inside Anthropic's Red Lines: The Girls' School That Doesn't Even Violate Them!

EPISODE 2 · THE STORY, WITH ITS SOURCES

*"Every coin has two sides, and we have always looked at both.
Today we look at the other one."*

The episode as a written story · every quotation referenced · 10 sources

ISSUE · SEPTEMBER 2026 · FACELESSYOUTUBEAUTOMATIONHQ.COM



AI GENERATED 
TRANSPARENCY DISCLOSURE

WHAT THIS DOCUMENT IS

The story of the episode, in the order it is told

This is the episode as a written story: what the video says, chapter by chapter, in the order it says it, condensed from the spoken text and adding nothing to it. Every quotation is reproduced word for word, as spoken or written by the person named, and carries a number in square brackets that points to the source list at the end. The reported figures carry one too. Where the episode says that something is reported rather than proven, this text says so as well.

00:00 · WHERE WE STAND

THE DISCLAIMER THAT OPENS EVERY EPISODE, WORD FOR WORD

Before we start, one thing about where we stand.

We are not against artificial intelligence.

We use it every day. Even this video is made with it.

This channel exists to make artificial intelligence easier to work with. To help people use it better, and with more purpose. And to take away the hesitation, so anyone can look at this technology and judge it for themselves.

But every coin has two sides, and we have always looked at both. Today we look at the other one.

Everything you will hear is sourced. Where something is reported rather than proven, we say so.

And artificial intelligence is not the only thing that can be wrong. So can we. We check every claim to the best of our knowledge — and where we are not certain, we say that too.

00:50 IN THE EPISODE

“It doesn’t even violate our red lines”

A journalist sits across from the chief executive of an artificial intelligence company worth close to a trillion dollars. Its stated mission is to build this technology safely. Its founders walked out of another company because they thought safety was losing.

She asks him about a missile strike on a girls’ school in Iran. More than a hundred and fifty people killed, most of them children. She asks whether his model played a role.

This is his answer: “We don’t know exactly how these models were used. Mistakes that happen in warfare are really, really terrible. And what you’re talking about is a use case that doesn’t even violate our red lines.”^[1]

Not: this must never happen. Not: we are investigating. It does not even reach the line.^[1]

01:49 IN THE EPISODE

What is proven, what is only reported

Everything in this episode comes from one place: a Bloomberg Originals documentary called The Circuit. It is public, and the link is at the end of this document. Some of what is in it are Bloomberg's own reporting claims.

Whether those reports are accurate, neither the channel nor the viewer can verify. What can be verified is that Bloomberg published them, put them to him directly, on camera, and he answered. That is the claim this episode makes.

Not that any particular strike involved his software. Nobody knows that, and he says so himself. What the episode shows is what a company says about itself when somebody asks the hard version of the question. ^[1]

02:39 IN THE EPISODE

The Pentagon contracts — \$200M each

The story starts in 2025. Anthropic, OpenAI, xAI and Google each win US Pentagon contracts worth up to two hundred million dollars, two-year deals for agentic AI workflows. Anthropic delivers through Claude Gov, for classified workloads.

For this company it is worth pausing on, because as a physics undergraduate at Caltech, Dario Amodei had a reputation as an anti-war advocate.

Emily Chang puts that to him: a longstanding anti-war stance, and yet one of the first AI companies to operate on the networks the United States uses to fight wars.

His answer: "Look, the world changes. When I see Russia invading Ukraine, when I see the risk of China invading Taiwan, it worries me that we have a kind of resurgent authoritarian block, and that we need to defend ourselves." ^[1]

That is a position a reasonable person can hold. Then Bloomberg reports that Claude was used by the US military in the operation to seize the Venezuelan president, Nicolás Maduro. And weeks later, everything came apart. ^{[1][2]}

04:02 IN THE EPISODE

The ultimatum, the blacklist, the judge

The Department of Defense demanded full use of the technology, without guardrails. Anthropic refused and drew two red lines: no mass surveillance, no autonomous weapons. The government's answer was a deadline, then a blacklisting, then lawsuits.

The Secretary of Defense, Pete Hegseth, on the record in a Senate hearing: "Anthropic is run by an ideological lunatic."^[5] Amodei, asked how that feels: "I've been called worse things than that all the time."^[1]

And asked what winning looks like: "I won't even call it a fight. This is more a debate about what the proper use of AI by the government is."^[1]

On March the twenty-fourth, US District Judge Rita F. Lin said from the bench that the ban looked like an attempt to cripple Anthropic, and that she was concerned the company was being punished for public criticism.

So far, this is a company taking real damage for a real position. The rest of the episode is about what that position actually covers. ^{[4][5]}

05:20 IN THE EPISODE

A thousand targets a day, then five thousand

Emily Chang brings a number into the room. A US government official has said that with the help of large language models, the military has gone from being able to hit a thousand targets a day to five thousand.

That means Claude can help kill more people, more quickly. Is he comfortable with that?

"Basically you're asking, do you believe in this country. Do you want this country to be a more powerful actor rather than a less powerful actor on the world stage. I do. I'm a patriot."^[1]

And then: "It's not up to me. If we provide a technology, it's not up to us to say, you can do this military operation and you can't do that military operation. Now, I might privately believe that this military operation makes sense and that military operation is a bad idea, but we're not gonna deny the technology. You have to leave policy in the hands of the military decision makers."^[1]

Read that against the two red lines. Everything on the other side of them is, in his own words, not up to him. ^{[6][7]}

06:25 IN THE EPISODE

The girls' school

Then she goes further. Bloomberg has reported that Claude is being used by the US military in the war in Iran, for AI-assisted targeting, through a platform built by Palantir called the Maven Smart System. In February, a US missile reportedly hit a girls' school in Iran.

More than a hundred and fifty people were killed, most of them children. Did Claude play a role in that strike?

"We don't know exactly how these models were used. Obviously mistakes that happen in warfare are really, really terrible. This is a really terrible thing to happen. We were willing to risk the future of our company to limit how these models are used — and what you're talking about is a use case that doesn't even violate our red lines. We're worried that there will be a hundred times as much with use cases that do violate our red lines."^[1]

And: "I would say overall the use of these models is appropriate. I think it's good on net. But military decision makers make terrible mistakes even at the best of times, and I don't know if we're in the best of times."^{[1] [8]}

07:24 IN THE EPISODE

His full answer — and the 69 seconds

He did not dodge, and he did not pretend to know more than he does. But look at what the sentence contains. He places a hundred and fifty dead, most of them children, in the category of things that do not reach his red line.

He does it calmly, in a full sentence, without being pushed. And the concern he expresses is about the cases that would reach it. There is a second thing in it, visible only next to something he said sixty-nine seconds earlier.

Sixty-nine seconds earlier, he handed the decision to the military decision makers. Then he tells us those same people make terrible mistakes even at the best of times, and that he is not sure these are the best of times. ^[1]

08:20 IN THE EPISODE

The follow-up question

Emily Chang does not let it go. This school had a website. It could have been found in a Google search. Shouldn't Claude have spotted that? And doesn't it speak to a scarier issue about using technology as a shortcut in war?

His answer: "The principle that was obeyed here is a human makes the final decision. I don't know what role Claude or any other AI had. But if this isn't an illustration of why that principle is so important, I don't know what is."^[1]

That is a real answer with a real argument in it: a human made the call, the machine assisted. It also means this. The safeguard he is pointing to is not in his product.

It is in somebody else's chain of command, the same chain of command he had called fallible one minute before. ^[1]

09:14 IN THE EPISODE

Project Glasswing and the NSA

Now the part that closes the circle. While Anthropic was blacklisted by the Department of Defense, it was also running Project Glasswing, the program that gave a small number of selected organizations access to Mythos, the version of its most powerful model with the safeguards switched off.

Bloomberg reports who wanted in: US federal agencies, including the National Security Agency.

So at the same moment: banned by one arm of the US government for refusing surveillance and autonomous weapons, and handing out the unrestricted model to another arm of the same government, under a vetting program the company itself designed and controlled.

Asked why anyone should be comfortable with that concentration of power, Daniela Amodei answered: "It wasn't like, oh, it's so powerful and let's decide who gets the power. It was a very specific concern around cybersecurity. There's obviously nuance to decide where you draw that circle. We've tried to be as publicly open as possible to say we're trying our best to make this decision well, but we might not do it perfectly."^[1]

They decide who gets the strongest model in the world. They wrote the criteria. They apply them.^[1]

10:28 IN THE EPISODE

What he gets right

To be fair to him, because this is the part the outrage clips leave out, he says things in this interview that most executives would never say on camera. He says he is scared of companies having this technology, and scared of governments having it.

He calls the current arrangement dangerous and unstable. He calls for required pre-release testing and auditing of models, by law. And he puts a number on the risk: a ten to twenty-five percent chance of civilization collapse.

Emily Chang holds it up against him: nobody would board a plane with a twenty-five percent chance of crashing. "That's right. Twenty-five percent is too high."^[1] He knows. He says it out loud. And then the next model ships.^[1]

11:18 IN THE EPISODE

What this video does not claim

There is an easy version of this story, and the episode does not make it. The easy version says Anthropic's AI killed those children. That is not established, he does not claim to know, and the episode does not say it.

What is established is narrower, and worse. Dario Amodei is not a bystander. He is the co-founder, the chief executive and a part-owner of the company that supplied the artificial intelligence to the system that does the targeting in that war.

That is not an accusation; it is his job title and the reporting, and none of it is in dispute.

He was Vice President of Research at OpenAI for about five years, co-built GPT-2 and GPT-3, left in December 2020, and in 2021 founded Anthropic with his sister Daniela. The reporting on why has been consistent ever since, and it comes down to two words. ^[1]

12:27 IN THE EPISODE

The first word: commerce

The first word is commerce. OpenAI had turned commercial: a billion dollars from Microsoft, a non-profit restructured into a capped-profit company. Six years later, in February of this year, Anthropic raised thirty billion dollars at a valuation of three hundred and eighty billion.

In May it raised sixty-five billion more, at nine hundred and sixty-five billion. Its revenue run rate crossed forty-seven billion. On the first of June it filed to go public, holding a Department of Defense contract worth up to two hundred million dollars.

OpenAI's last round, in March, valued OpenAI at eight hundred and fifty-two billion. The company he walked out of because it had gone too commercial is today worth less than the one he built to correct it. ^{[3][9][10]}

13:45 IN THE EPISODE

The second word: safety

The second word is safety. Bloomberg asked him about a strike on a girls' school that killed more than a hundred and fifty people, most of them children, in a war where his model is reported to be running inside the targeting platform.

And the answer was that this is a use case that does not even violate their red lines. That is not the register of a man describing a catastrophe. It is the register of a man describing a category.

There is no published threshold anywhere. There is no number. There is no rule in that company that says: at this point, we stop. There are two prohibitions, mass surveillance and autonomous weapons, and beyond them, in his own words, it is not up to him.

Almost every sentence used to state that position is his. ^{[1][8]}

14:54 IN THE EPISODE

The question

A man leaves the company he helped build over two things: too much commerce, too little safety. He founds the one that is going to do it properly.

Six years later his company is valued at nine hundred and sixty-five billion dollars, more than the one he left, and a hundred and fifty dead children do not reach its red line. Both reasons he left, now on his side of the table.

Is that a defensible position for a safety company to hold? Or is a red line with no number behind it simply a line you can always say you have not crossed? And if you think there should be a number, what is it?

Because that is the one thing nobody in this industry has been willing to write down. ^[9]

SOURCES

Every claim, with its source next to it

10 sources, in the order the episode uses them. The bracketed numbers in the text lead here. Timecodes match the chapter marks on YouTube.

[1] Inside Anthropic: The Circuit, with Emily Chang

Bloomberg Originals · 2026

The on-camera interview. Every quotation in the episode comes from this recording; where Bloomberg reports something rather than shows it, the episode says so.

<https://www.youtube.com/watch?v=v1wZwxY3CMg>

[2] Anthropic, Google, OpenAI and xAI granted up to \$200 million from the DoD

CNBC · 14 July 2025

The contracts themselves: two-year deals for agentic AI workflows.

<https://www.cnn.com/2025/07/14/anthropic-google-openai-xai-granted-up-to-200-million-from-dod.html>

[3] Anthropic CEO says why he quit OpenAI

Yahoo Finance

Why he left OpenAI, in his own words.

<https://finance.yahoo.com/news/anthropic-ceo-says-why-quit-194409797.html>

[4] Judge on Anthropic's Pentagon injunction hearing

NPR · 24 March 2026

US District Judge Rita F. Lin from the bench — the line quoted in the episode.

<https://www.npr.org/2026/03/24/nx-s1-5759276/anthropic-pentagon-claude-preliminary-injunction-hearing>

[5] Hegseth calls Amodei a 'lunatic' and defends Pentagon use of AI

Bloomberg · Senate hearing, 30 April 2026

Pete Hegseth, U.S. Secretary of Defense, in a Senate hearing. Bloomberg's page is behind a paywall; title and date as published.

No public link — title and date as published.

- [6] Iran war tests Project Maven and the U.S. AI war strategy**
Bloomberg Businessweek · 12 March 2026
The targeting numbers, and the report that Claude runs inside the Maven platform. Reported, not shown.
<https://www.bloomberg.com/news/features/2026-03-12/iran-war-tests-project-maven-us-ai-war-strategy>
- [7] What is the Maven Smart System and what does it do?**
CSIS
What the platform is, explained by a think tank rather than by a vendor.
<https://www.csis.org/analysis/what-maven-smart-system-and-what-does-it-do>
- [8] Deadly Iran school strike casts shadow over Pentagon's AI targeting push**
Military Times · 24 March 2026
The strike, the casualty figures, and the open question of what role AI targeting played. Nobody has established that Anthropic's model was involved.
<https://www.militarytimes.com/news/your-military/2026/03/24/deadly-iran-school-strike-casts-shadow-over-pentagons-ai-targeting-push/>
- [9] Anthropic raises \$65B Series H at \$965B post-money valuation**
Anthropic · 28 May 2026
The valuation, in the company's own words.
<https://www.anthropic.com/news/series-h>
- [10] Anthropic raises \$65 billion, nears \$1T valuation ahead of IPO**
TechCrunch · 28 May 2026
The round, the IPO filing, and OpenAI at \$852bn.
<https://techcrunch.com/2026/05/28/anthropic-raises-65-billion-nears-1t-valuation-ahead-of-ipo/>

ON THE EVIDENCE · CORRECTIONS

Not a claim that Anthropic's model played any role in the strike on the school. Nobody has established that; Dario Amodei says himself he does not know. Not a claim that anyone broke the law. Short third-party excerpts are used for criticism, commentary, news reporting and analysis; all rights remain with their owners, credited above. If anything here is inaccurate, tell us and we will correct it — in a pinned comment on the video and, where it matters, in this document. That standard applies to us the same way it applies to everyone else in this episode. Write to us at facelessyoutubeautomationhq.com/contact.html.

WATCH THE EPISODE · EVERY SOURCE, ONLINE

The Other Side · Episode 2

Inside Anthropic's Red Lines: The Girls' School That Doesn't Even Violate Them! — 15:52, live since 15 September 2026. This channel is made with artificial intelligence, and it says so in every frame. The same sources as above, with live links, are on the resources page of the channel.

youtu.be/sJFtY_JL6cE →

facelessyoutubeautomationhq.com/resources/the-other-side-02-inside-anthropics-red-lines/ →