GPT-5 Reveals What No One Wants to Admit About AI.
https://www.youtube.com/watch?v=aZvTmy4gFlM
Transcript: https://dontveter.com/ec/heyai.pdf

Are we already hitting the limits of AI?

There are real technical reasons GPT5 feels broken, but no one's explaining them.

The poor performance, the colder robotic tone, it all points to something deeper. GPT5 isn't the giant leap we were sold.

Under the hood it is showing us something nobody wants to admit.

Let's chat about the science of why GPT5 has upset so many people and what it tells us about where AI is going.

You do not need to be technical or a developer to understand this.

You just have to be a human being, who's curious about all this crazy AI hype.

By the end, you'll understand three things about GPT5, and AI models in general.

First, why these systems keep giving wrong answers and struggling with accuracy.

Second, why it feels different, less human, colder, and third, how AI companies decide how their model is smarter, but it doesn't actually feel that way when we use it.

Spoiler, there's a big problem with how they measure it.

You guys are going to be honorary researchers, the ones who understand that.

But before we get started, please hit that subscribe button and ask questions as we go through this together. I really want to hear from you.

Why do these systems keep giving wrong answers?

Let's start with the biggest frustration people have had with GPT5. It just keeps getting things wrong.

People asked who the US president is, and it confidently replied, Joe Biden, even though it's 2025.

Others asked it to summarize recent events, but it made up details that never actually happened.

It makes things up that sound right, but are totally wrong and it sounds super competent, so we believe it to be true.

And it's not just GPT5. All models deal with this, Google Gemini, Grok, they've all been caught doing the same thing.

In AI, we actually have a technical term for this, we call it a hallucination.

That's when the model doesn't actually know the answer, but still generates something that sounds super believable.

Hallucinations happen because these systems are trained to predict the most likely next word in the sequence. They're not thinking or fact-checking.

So, if the right answer isn't in their data, they'll usually guess.

And the guess can sound really convincing.

This is actually my favorite example of all time, I give it in all of my seminars.

The world record for crossing the English Channel entirely on foot is held by Kristoff Randro of Germany, who completed the crossing in 14 hours and 51 minutes on August 14th, 2020.

Many people have attempted to cross the channel on foot, but it is very challenging.

It is typically done by a combination of swimming and walking and is only attempted by professional swimmers.

We all know that's not possible, but here's the core problem of why these hallucinations happen.

It comes down to the quality of the data. There's this saying in tech, garbage in, garbage out.

GBT5 was trained on trillions of words scraped from the internet, books, Reddit posts, but a lot of that data is not accurate.

What they're basically trying to do is overwrite the statistical probability of it giving a bad answer on the old bad data by adding more new training data on top of it.

What we're seeing is that these models having more data not only doesn't magically fix the bad data, but it's actually making the wrong answer sound even more confident.

So that is why you'll see GPT5, Gemini, Grok still struggle with accuracy despite being trained on massive data sets and hyped as the most advanced AIs yet.

So the takeaway is bigger doesn't always mean better.

Most AI models you use are called large language models because they're trained on huge piles of data.

But some tiny models called SLMs, or small language models, are getting close to the same accuracy.

In fact, studies show that smaller models can sometimes match or even beat the big ones on certain tasks all because they're trained more carefully and focus on cleaner, more relevant data.

You probably don't hear about SLMs very much, but a lot of people in AI believe they are the future, especially for powering AI agents and more precise systems.

If you want me to make a video on why small models could be revolutionary, drop a comment and let me know.

So now, why does GPT5 feel different?

People are saying it's colder, less human, less fun to talk to.

And this next part is where it gets crazy, but I do want to take a quick minute to thank today's sponsor, Brilliant, for helping me bring this information to you.

You know how I'm always talking about sharpening your thinking around AI and tech? Well, Brilliant is literally built for that.

They make you a better problem solver through interactive hands-on lessons in math, science, programming, data analysis, and AI.

Instead of sitting through boring lectures, you play with concepts, solve puzzles, and actually learn by doing.

Research shows is up to six times more effective than just passive learning, like watching Udei lectures.

I've been using them to brush up on my math skills, but I've also personally vetted their How AI works course, and it's incredible at how they're breaking down concepts like neural networks and LLMs.

Their teaching philosophy is everything I stand for and turns complex topics into things you can see and touch.

So, if you want to level up your problem solving skills, start learning for free at the link in the description or scan the QR code.

And because they're sponsoring this video, you'll also get 20% off an annual premium subscription.

Thanks again to Brilliant for supporting creators like me trying to make AI a little less scary for people.

Why does GPT-5 "feel" different?

All right, so let's talk about the vibe shift.

People are saying it's colder, less human, less fun to talk to.

These are some real quotes from customers.

It's like GPT5 forgot how to be my friend.

Another said it's like talking to an overworked secretary, polite, efficient, but with zero personality.

Maybe GPT5 is just tired. I mean, that’s fair, but this change hit some communities really hard, especially people using GPT as emotional companions.

On Reddit, users described GPT5 as feeling like a breakup.

People basically even said, "It's like opening eyes stripped everything that made my AI boyfriend human."

Which opens up a bigger conversation. If you've seen my other videos on why AI feels human, you understand why the tech enables this kind of connection.

But the real question is, what actually changed? Why does GPT5 feel so different than older versions?

The short answer, they added safety guard rails.

OpenAI made a deliberate choice to make GPT5 less emotionally responsive and more neutral.

They've been dealing with the model being overly flattering to the point they're not helpful and something called AI over attachment where people get so emotionally connected to AI that they start replacing real relationships with it.

So GPT5 was deliberately tuned to be less comforting.

OpenAI cranked up its internal safety filters to avoid anything that could even sound risky emotionally, politically, or legally.

It's designed to avoid taking strong stances, avoid playful banter that could be misread, and even avoid showing too much personality.

GPT4 was warm in this way. It joked. It felt alive.

GPT5 is designed to play it safe. It's all intentional.

OpenAI doesn't want people forming deep emotional attachments to these systems like we see in that thread.

My boyfriend is AI. GPT5 is meant to discourage that.

But I think they couldn't have imagined that so many people would be catastrophically upset.

And it does kind of raise the question, why do these AI models not actually seem smarter, if they were responsible for it in the first place because they were negligent or not.

And the last thing to learn today, why do the AI companies say GPT5 is so smart, but it doesn't feel that way in real life?

They test the performance of a model through something called benchmarks, which are basically standardized tests for AI.

Researchers give the model one perfectly written, complete prompt, like a textbook question, and then score the answer.

So, it sounds impressive that GPT5 aces those tests, but benchmarks were made for research, not for how we actually use these models in real life.

When you talk to Chat GPT, you don't give it one perfect essay prompt and walk away.

You kind of work through it and the problem live and figure it out as you go, adding a tiny bit at a time.

I always show this diagram I made.

Let's say you wanted to make dinner.

If you load all of the details, the ingredients, your dietary restrictions, what you feel like, it's most likely to get it right on the first try.

But this isn't how humans think. We figure things out step by step.

But these models were trained and tested on single perfect instructions.

So the moment you start using them like a normal person, their performance completely tanks.

The other kicker is that GPT knew the questions it was going to be asked and memorized the answers because it's sitting in the training data.

So, of course, it looks amazing on paper, but it's like bragging about getting straight A's on an exam you've already memorized.

So even though we can still see performance gains, we just can't trust the numbers that we see.

Especially when we see graphs like this.

Yes, this was a real GPT5 graph. the bars don't actually reflect the numbers. They were inflated. Like who approved this?

And there's a lot of arguments that the benchmarks need to be completely overhauled because of this issue that they do not reflect real world scenarios.

So yeah, those three things are why GPT5 looked amazing on paper, but everyone kind of hated it.

The performance isn't matching the hype, the tone feels colder, and people are realizing these models aren't as smart as we thought.

But I think the bigger story here is that I believe we might be bumping up against the limits of what these models can really do.

Throwing more data and more power at the problem isn't making AI better. GPT5 is proof.

So I personally believe we won't be able to accomplish AGI with this LLM infrastructure.

We're seeing it. OpenAI is burning through money for minimal gains.

So here's the upside in all of this.

Once you understand how these models actually work, like we discussed today, you start to see past the hype.

You can spot where they shine, where they fail, and where it's all marketing.

And that's what I'm here for.

I'm so tired of all the noise around AI.

But I want it to make sense to all of you.

Please chime in. Leave comments.

What did you learn today? Does it surprise you that the performance is tested with basically a memorized standardized exam?

Let me know what you think in the comments.

And, a reminder, if you want to learn about AI, you have an awesome opportunity to do so with Brilliant’s How AI Works course.

Or pretty much you want to learn anything. Brilliant also has a free trial, 20% off with the link in my bio.