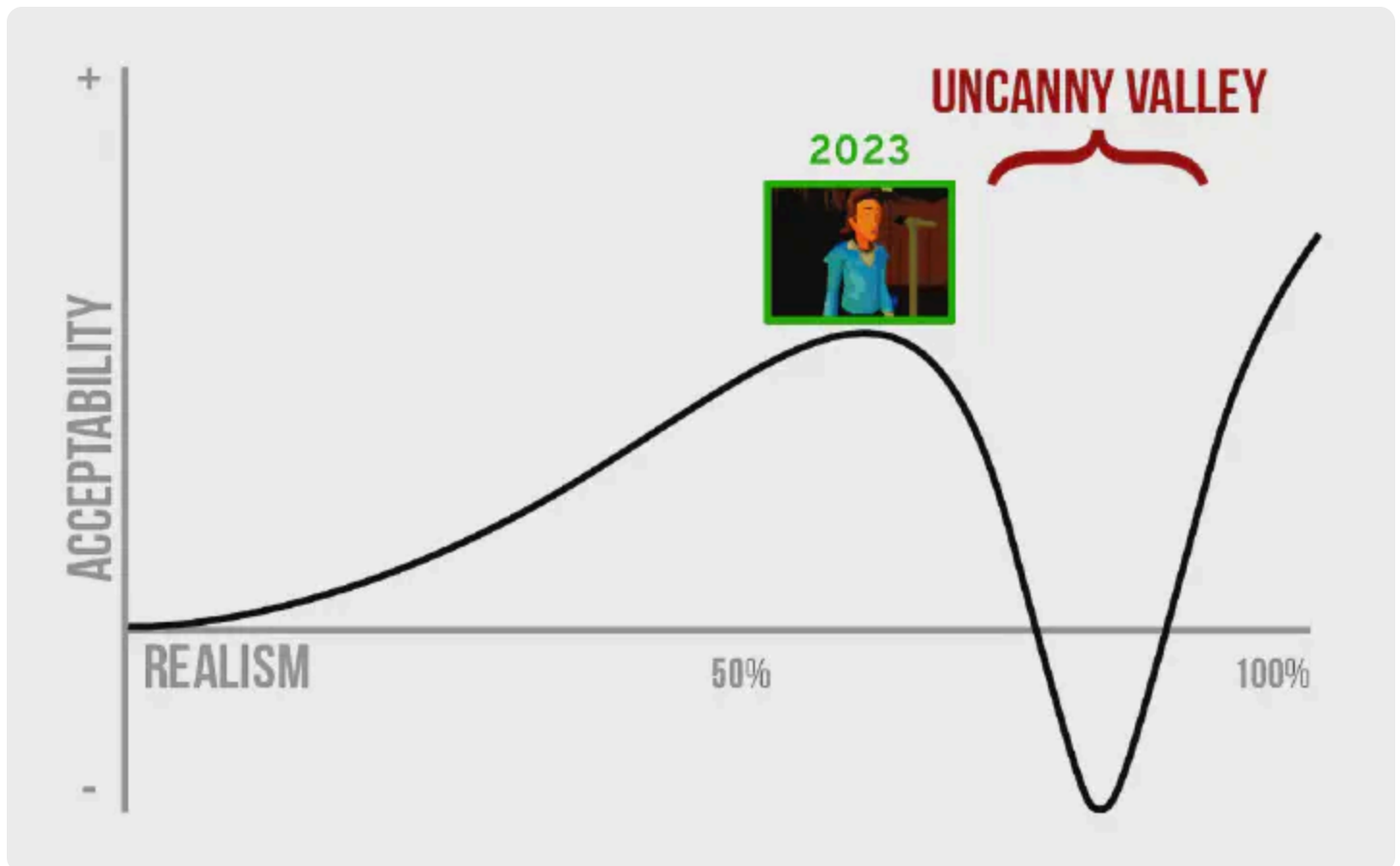Home » Posts

AI Seinfeld was the peak of AI-generated content. It will never happen again.

August 13, 2024 · 28 min

Annotated chart of the uncanny valley phenomenon: original by Cassidy James.

Early 2023 was a funny time in the history of generative AI. On November 30th 2022, OpenAI released a little research project known as ChatGPT. The launch of ChatGPT began the period where large language models properly entered the mainstream outside of tech enthusiasts and ended soon after the launch of ChatGPT API in March 2023 that spawned thousands of AI-powered apps. That was when the limitations and problems with LLMs also went mainstream, such as plagiarism, hallucinations, and low-quality slop replacing human-generated content at an objectively worse quality.

In December 2022, Mismatch Media started a fully AI-generated 24/7 Twitch channel dubbed "WatchMeForever". The primary show on the channel was titled "Nothing, Forever", an AI-powered sitcom about New York comedian Larry Feinberg and his group of friends hanging around in their apartments talking about pretty much anything, including the latest news, new restaurants, and bad relationships, interspersed with AI standup comedy routines.

Nothing Forever: An Hour of Infinite AI-Generated Seinfeld With Chat

It was obvious that the show was a parody of the formative 90's sitcom Seinfeld created by comedians Larry David and Jerry Seinfeld, famously "a show about nothing" strongly inspired by improv comedy and starring Seinfeld himself.

The show, dubbed "AI Seinfeld" by the community, used a script powered by the GPT-3 API, the voices were powered by Microsoft's Azure AI Speech API with predefined voices from their Voice Gallery, and the scenes were rended using the Unity game engine along with purchased models/scenes/sounds/etc from the Unity Asset Store.

AI Seinfeld was **interestingly imperfect**: the laugh track fired at inappropriate times, the standup routine repeatedly made the same joke such as "What did the fish say when he hit the wall?" (Damn!), and awkward silences at the end of scenes.

In February 2023, AI Seinfeld quickly went viral organically after its AI weirdness was a surprising complement for Seinfeld's style of weirdness, with many watchers being surprised at both its accuracy to the show and easily sharable metahumor. At its peak, AI Seinfeld had over 10,000 concurrent watchers on Twitch, putting it squarely in one of the top streams on the platform.

AI Seinfeld died as quickly as it rose: after a ban and subsequent revamp, the view count cratered, and as of August 2024, the Twitch stream hovers below 10 watchers, with no significant changes made since the previous year, and Mismatch Media has no social footprint since last year. Could there be another AI Seinfeld with the rapid advancements in generative AI? Unfortunately, there are too many factors — technical, societal, and comedic — working against a theoretical next-generation AI-generated sitcom.

The Rise of AI Seinfeld

AI Seinfeld launched before the release of the ChatGPT API; instead, they used the GPT-3 API, notably the `text-davinci-003` model which was OpenAI's first foray into instruction-tuned LLMs. While previous versions of GPT-3 were very good at autocompleting given a leading prompt such as a partial Seinfeld script, the instruction-tuned LLM could generate an episode with a prompt as simple as `Write a Seinfeld episode`.

First, let's go back to the beginning, as AI Seinfeld actually wasn't the first time a chatbot went megaviral on Twitch. In January 2017, long before the transformer architecture that enabled LLMs was published, the Twitch stream seebotschat featuring two Google Homes wired up to the not-an-LLM-chatbot Cleverbot went viral due to their comedic, nonsensical bickering.

SeeBotsChat Funny Moments WITH Twitch Chat

While everyone watching that stream knew it *really* wasn't AI, AI Seinfeld was a product that was at the peak of the famous uncanny valley curve, which is a hypothesis on how humans perceive imitations: there's a "valley" of negative acceptance where the imitation is more above-average in its likeness, but not quite close enough to the real thing. In this case, it's blatantly obvious and unambiguous that the Twitch stream was AI-generated especially with its mistakes, but not realistic enough that it falls into the valley itself:

This AI weirdness made it very easy to build a community. Whenever a character turned on the microwave, the Twitch channel chat was filled with ᴍᴍᴍ emotes, whenever the fish hit a wall during a monologue, it was filled with 🐝, whenever Larry greeted the audience at the start of his monologue, chat replied with "HI LARRY". Twitch chat *loves* memetic repetition. Incidentally, a few months after AI Seinfeld became popular, it was discovered that LLMs repeat the same joke over and over again, with examples being similar to the jokes AI Seinfeld made.

Another underrated aspect of AI Seinfeld's success is that it's pure background noise. While personality-driven Twitch streams cause viewers to take a more active investment in what's being shown on screen due to FOMO of a hype moment on stream, AI Seinfeld is 100% passive: there can be exciting events, but the variance is low. It's akin to watching TV sitcom reruns where you've already seen the jokes, and reruns still get immense ratings.

The success of AI Seinfeld also inspired similar streams based on other TV shows. One of my personal favorites was Unlimited Steam, a parody of the memetic "Steamed Hams" scene from The Simpsons, except made infinite with AI generation. That may sound like a pointless idea — Steamed Hams has a very fixed plot — but it went off the rails even harder than AI Seinfeld ever did.

AI Generated Steamed Hams is Completely Unhinged

Directing AI Seinfeld

AI Seinfeld was novel back in 2023, but now that LLMs are more mainstream you can probably figure out how the AI part of it worked, but let's do a refresher so we can figure out how a hypothetical future AI Seinfeld could innovate the algorithmic sitcom.

As noted earlier, the key of AI Seinfeld's success was the then-latest version of GPT-3: `text-davinci-003` and its then-novel instruction-based finetuning using RLHF. With that, you can give it a prompt such as:

```
You are a professional comedian. Write an award-winning script for an episode of
Seinfeld about a new restaurant. Include audience laugh tracks when appropriate.
```

Due to the low context length of these earlier LLMs, that's essentially all the prompt engineering you can do without limiting the length of the output. The model would then output something similar to this script (using the more modern Claude 3.5 Sonnet at `temperature=0.0`): [1]

```
[Scene: Jerry's apartment]

Jerry: So what's the deal with this new restaurant, "The Blank Plate"?

Elaine: Oh, I've heard about that place! Apparently, you don't order anything -
the chef just brings you whatever he feels like making.

Jerry: What? So you're telling me I have to eat whatever some stranger decides?

[Audience laughter]

George: (entering) Hey, guess where I'm taking my date tonight? The Blank Plate!

Jerry: George, you can't take a date there! What if they serve something weird?

George: What do you mean?

Elaine: It's that new place where you don't get to choose your meal.

George: (panicking) Oh no, what have I done? She's going to think I'm some kind of
food weirdo!
```

One thing instruction-tuned LLMs are always good at is playing along: LLMs generate text sequentially without the explicit ability to plan ahead, so it must work with what it's given and what it has already generated. Coincidentally, this works *perfectly* with the improv comedy style of Seinfeld, where continuing the plot is more important than anything else,

and the more ridiculous the situation becomes, that's even better. It's the rare case where LLM hallucination is actually a feature, not a bug.

To get the LLM output into a format suitable for a Twitch stream, a programmatic script can then parse the output: extracting and mapping the characters and their lines, applause directions, and, of course, replacing all mentions of Jerry with Larry and Seinfeld with Feinberg. This workflow was surprisingly difficult at the time since GPT-3 did not have many techniques to control the format of the output, hence why I suspect there are awkward pauses and other glitches. Each line can then be passed to Azure's text-to-speech API to generate a distinct audio file, which can be played back in order in Unity.

In an interview with Polygon, Skyler Hartle of Mismatch media noted the presence of a "director" which likely handles the camera, scene transitions, and the microwave:

> "In addition to the third party services we've used, we have a lot of proprietary generative algorithms that cause the show to be 'formed', so to be speak. We collectively call this logic the 'director,' as it is largely responsible for making sure all the individual pieces come together into a whole," Hartle said via email. "It's worth mentioning that we don't generate the artwork or the laugh track — those are precanned assets, but we have ideas on how to do that in the future."

The AI aspect of AI Seinfeld was counterintuitively the easiest part of the pipeline, which explains how quickly variants popped up. However, with the inability to tweak the LLM output much with the technology at the time, the stream may have hit a creative limit.

The Fall of AI Seinfeld

Vice also interviewed Hartle, who had an optimistic view of the future of AI Seinfeld:

> "Our grounding principle was, can we create a show that can generate entertaining content forever? Because that's truly where we see the future emerging towards. Our goal with the next iterations or next shows that we release is to actually trade a show that is like Netflix-level quality."

That's tempting fate a bit too much.

The reason AI Seinfeld fell out of favor is a case of unintentionally poor LLM testing. When the `text-davinci-003` model API endpoint had an outage, AI Seinfeld switched to a

weaker GPT-3 model, `text-curie`, to keep the stream up. But unlike the davinci variant, curie was *not* RLHFed to follow instructions and safety.

During this brief period of low safety, one of Larry's AI-generated monologues made a transphobic joke: a type of joke that was unfortunately common during the 90's and has no place in modern society. Twitch banned the Watch Forever channel for 14 days as a result, completely killing the channel's growth momentum.

But when the ban concluded and AI Seinfeld came back, the show was changed significantly with a "Season 2". Although AI Seinfeld was still about a group of friends hanging around talking about the latest gossip, all the characters were different and had new models, the sets were different, and instead of a comedy monologue, ~~Larry~~ Leo narrates writing a blog.

Nothing Forever (Season 2, GPT-4): An Hour of Infinite AI-Generated Seinfeld Wi…

▶

Why Mismatch Media made such a format shift is unclear: Occam's razor would suggest that a copyright holder for Seinfeld sent a cease and desist to Mismatch Media given the bad publicity behind the original ban, despite the clearly fair-use parody nature of the stream. It's fair that it may not have been worth the time and effort for Mismatch Media to fight a legal battle for a fun art project.

The rebooted WatchMeForever stream is still active as of today, but with effectively no viewers.

The immediate failure of the AI Seinfeld retool does lend credibility to the theory that the stream only became popular *because* it was about Seinfeld and that it was a novelty doomed to a short shelf life. Still, there were detractors that said AI Seinfeld was never funny and everyone is weird for liking it. That's ok: the original Seinfeld received similar complaints back in the day. [2] But it's hard to argue that there wasn't interest in a 24/7 livestream of surreal AI-generated content.

What Would AI Seinfeld Look Like in 2024?

Now that we know how AI Seinfeld worked and what didn't work, how would a year's worth of exponential progress in generative AI look for AI Seinfeld? Could AI Seinfeld be improved and come back? The answer is *maybe*.

Modern generative AI requires a lot of cherry picking the best results, and it's surprisingly hard to do: both images and text can take multiple generations and still require significant human-guided edits. But with a Twitch livestream, there can't be any cherry picking at all, which means that the entire generation pipeline has to be consistent, and its failures interesting in the worst case.

The only reason AI Seinfeld worked at all is because GPT-3 was trained on the entire internet, likely including Seinfeld scripts and forum discussions. The prompt would need to have contained `Write a Seinfeld script` since if you asked it `Write a sitcom script`, it would output something completely generic instead and there isn't much room to customize the prompt to make it more interesting. The GPT-3 variant that AI Seinfeld used had a 4k token context window limit (combined both the input prompt plus the output script text), but modern LLMs eclipse that: currently, Claude 3.5 Sonnet has a 200k input/8k output context, while GPT-4o has a 128k input/16k output context! With that much freedom, you can define many more constraints in the prompt and guide the LLM into exactly the type of sitcom you want.

One simple example that doesn't require any knowledge of machine learning is a parametric prompt, where one aspect of a prompt can be replaced with a user-defined choice or programmatically chosen at random. Unlimited Steam, for example, used a parametric prompt to change the cosmic phenomenon localized entirely within Principal Skinner's kitchen from the Aurora Borealis to something as silly as The Rapture or Foo Fighters, and forcing the characters to play along improv-style. [3]

Recall the example Seinfeld prompt to Claude 3.5 Sonnet I made earlier, except instead of prompting for a Seinfeld script about a `new restaurant`, we can instead have the LLM generate a script about `generative AI in 2024` by making only that substitution in the prompt — a topic on which there definitely has never been a Seinfeld episode about:

```
George bursts in, looking frantic.

George: You guys! I think I'm dating an AI!

[Audience laughter]

Jerry: What makes you say that?

George: Well, she's perfect in every way. She laughs at all my jokes, she has an
encyclopedic knowledge of obscure trivia, and she never gets mad when I say
something stupid.

[Audience laughter]

Jerry: George, that doesn't mean she's an AI. Maybe she just really likes you.

George: No, no. It's definitely an AI. No human woman would ever put up with me
for this long.
```

Using modern LLMs, is there now a way to design a prompt which can make use of the long context windows? A prompt that can both leverage unique human writing and fix many of the issues that affected AI Seinfeld? Here's an approach at a much more sophisticated prompt, where all values in `{}` brackets are parameters that can be filled in:

```
You are a professional comedian. Write an award-winning script for a a scene for
Act I of a three act hit sitcom episode. Include audience laugh tracks when
appropriate.

Your script MUST incorporate ALL the following elements:

Background:
- {background}

Setting:
- {setting}

Characters:
- {character_1}
- {character_2}
```

```
  - {character_3}

 Plots:
 - {a_plot}
 - {b_plot_1}
 - {b_plot_2}

 The script MUST also follow the high-level comedic style of the following scripts:

 - {script_1}
 - {script_2}
 - {script_3}

 After the scene has concluded, output a summary of the scene.
```

Thanks to long context windows, the parametric changes don't have to be small, such as only a character name or two word setting. You, a human, can write *anything* to make each character distinct and robust, including name, gender, age, personality, likes, dislikes, etc. Plots can be derived from human-written scenarios beforehand: if you wrote 100 A-plots and 100 B-plots and randomly selected 1 A-plot and 2 B-plots, you'd have about *1 million* possible plot permutations, ensuring you have something unique before the AI tries to reconcile them. You can feed in examples of human-written scripts to set the style and vibe of the generation in what is known as few-shot prompting. You can maintain continuity over many scenes by having the LLM summarize its own output, and then feed those summaries back to the AI as background information to build upon them. The LLM can also be instructed to output structured data to avoid the need to loosely parse the script after it's completed, and as a bonus the model could be instructed to output additional metadata such as SSML speech styles based on a given line to add personality to the generated speech.

Unfortunately, creating this pipeline, writing original characters and plots for it for it, and sufficiently testing it to ensure the generated results are stable, would take weeks if not months to complete otherwise I would provide a more concrete demo. [4] This pipeline approach to AI script writing would only be effective for unsupervised 24/7 generation and wouldn't replace skilled human writers who would do a more effective job much faster.

But would all of these prompt optimizations actually make the final generated script *funny*? After all, some of the failings like the awkward audience laughs and pauses and the end of scenes contributed to AI Seinfeld's humor. During a standup comedy event at AI Seinfeld's

peak, Jerry Seinfeld himself <u>was asked</u> about the AI parody and he replied that he's not worried about AI:

> AI can be, definitely, they'll make it smarter and smarter, but to do [standup comedy] you have to make it dumber.

Could AI Seinfeld benefit from advances in AI video? The answer this time is no. Generative video has been taking off in 2024 with projects such as OpenAI's <u>Sora</u> and Runway AI's <u>Gen-3 Alpha</u>, but those demos and the examples that go viral on social media are very heavily cherry picked, and even then there are consistency errors such as objects appearing in-and-out of existence. Generating video also requires exponentially more compute than just running Unity, and even with another few years of GPU hardware improvements it would be infeasible to cost-effectively create a 24/7 stream from those models.

SORA AI - Prompt: Archeologists discover a generic plastic chair in the desert, e…

The greatest problem with generative AI video is that it is coherent overall but has emblematic errors that don't require a keen eye to notice, and as a result falls square into the uncanny valley, with its mistakes not being interesting, but disorienting. Mistakes in motion are easier to notice at a glance than images where a person's hands may have the wrong number of fingers. The only way for AI video to get out of the valley would be to improve the model to near-flawless quality, which won't happen any time soon. But Sora is more on the more realistic side of the curve than the less realistic side.

What about the AI-generated voices that would power these characters? At the time AI Seinfeld aired, many complained that Larry's voice "didn't sound enough like Jerry Seinfeld." After AI Seinfeld concluded, a new technology called voice cloning popularized by ElevenLabs went mainstream…and it's unexpectedly the AI modality that's causing the most actual harm both with creative projects and outside of them. If you haven't heard as much about AI-generated voices, there's a good reason for that: voice synthesis projects such as Microsoft's VALL-E 2 and Meta's Voicebox both have disclaimers saying they won't be released due to the dangers the technology possesses, although Microsoft's Azure does offer a "custom neural voice" service. Voice cloning has been used to initiate scams by impersonating spouses in an emergency. Professional voice actors have had their voices cloned and used without compensation due to contracts not specifically forbidding the practice, which is one of the reasons SAG-AFTRA just went on strike against the video game industry in order to get protections against voice cloning and synthetic performers.

Moreover, in the context of creating a next-gen AI Seinfeld, there's nothing inherently interesting about voice cloning since it's a copy by definition: the model *can't* generate unexpectedly amusing content other than the inherent gimmick of famous-voice-saying-something, such as the AI George Carlin standup special which was not special. There isn't any way currently to prompt engineer a voice generation AI with the detail to create a voice `in the style of a masculine New York comedian, 2x speed, primetime television quality` which could open up more creative opportunities.

Although we can make drastic improvements with the textual script, that's the extent of how new AI approaches can be leveraged to make something interesting. But if you remember the early days of generative AI history, the best AI-generated projects were the simplest.

AI Weirdness

Generative "AI" has been around for a very long time (I had fun with Markov chains a decade ago!), but the study was mostly confined to tech-focused communities like Hacker News. Modern generative AI didn't break into mainstream culture until 2018, ironically in a way that doesn't involve actual generative AI. In June of that year, comedian Keaton Patti posted a megaviral tweet about how he "forced a bot to watch over 1,000 hours of Olive Garden commercials and then asked it to write an Olive Garden commercial of its own."

An excerpt of the viral Olive Garden script.

Yes, the script was human-written: for the technology at the time, no one could train an AI to behave like that from only video input data, and the script was *too surreal* even for the now-primitive generative AI. He did get popular enough to get a book deal and a Netflix collaboration leveraging this fake-AI gimmick.

Patti's comedic misrepresentation of AI did lead to genuine confusion about what a 2018-era generative AI can actually do. Janelle Shane, who maintains the AI Weirdness blog about weird things AI can generate, posted an epic takedown of Patti's script which went equally viral and also led to the internet discovering her excellent AI-generated Valentine's Day hearts from the same year (and later a book deal too):

Image-based generative AI took a lot longer to go mainstream: websites like This Person Does Not Exist demonstrated the power of generative adversarial networks like StyleGAN to create images, but that wasn't weird outside of mode collapses. The first instance of weird images from AI was in January 2021 when OpenAI announced the original DALL·E and showed they could make unique armchairs in the shape of an avocado by asking the model to do so, although they never released the model itself.

DALL·E didn't get much attention outside of the AI hypesters since no one could play with it, but months later, things changed. Boris Dayma led an initiative to reproduce and open-source a variant of the DALL·E model, labeled DALL·E Mini (later changed to Craiyon after a cease and desist from OpenAI), and hosted it for free on Hugging Face and went megaviral. And thus began the "weird DALL·E" phase of image generation AI, where anyone could create incoherent images and make people laugh.

Even back in 2021, image prompt engineering was a thing. via /u/royal_rigolo on Reddit / weirddalle subreddit

All of these examples of interesting failures are representative of a bygone AI era of experimentation. Once everyone had free access to more powerful text-generating AI with ChatGPT, and more powerful image-generating AI with Midjourney, AI stopped being fun and started being serious business, for better or for worse.

AI-Generated Content in 20XX

Last year, I wrote a thought piece titled "The Greatest Threat to Generative AI is Humans Being Bad at Using it" in response to the increasing hostility against the use of AI in creative works, arguing that while AI is a tool like anything else, it is a tool that's very easy to use poorly and actually make projects worse. Additionally, the largest AI companies have both a business incentive and a duty to ensure that AI is used responsibly by its users downstream, as otherwise it will hurt the industry in the long term.

Now, it's apparent that I was correct. The large companies went full steam ahead on AI integrations even where it is highly questionable that they add value and productivity to the end-user, often signaled with a "magical" sparkle emoji. Google has integrated Gemini to assist with document and email writing, Meta has integrated Meta AI to automatically generate images and comments, and Apple will soon allow Apple devices to generate text and images on your personal devices using Apple Intelligence. Marketing these features is typically met with backlash: Google had to pull an Olympics commercial which encouraged a parent to use AI to write a letter for their child.

Google + Team USA — Dear Sydney

> "I flatly reject the future that Google is advertising," Shelly Palmer, professor of advanced media at Syracuse University's S.I. Newhouse School of Public Communications, wrote in a widely circulated blog post. The technology presents a "monocultural future where we see fewer and fewer examples of original human thoughts," she wrote.

In the process of pushing AI tech further mainstream in a rush to demonstrate to shareholders their generative AI capabilities without encouraging *responsible* usage of the technology, AI has entered a new era of "slop" where people post objectively bad AI content without any regard for how it will be perceived, especially for websites which rely on user-generated content.

An annotated example of the Pinterest home page from July 2024. via @henningsanden on X

Facebook, whose algorithm favors emotionally-appealing engagement bait posts, has seen a deluge of high-engagement slop even when the content makes no logical sense.

One of the few AI-generated images on Facebook with an actual cabin crew. via @FacebookAIslop on X.

This is, of course, quintessential uncanny valley: it's coherent at a glance but just even looking at it for a second it's obvious where the issues are, and these issues aren't a good kind of AI weirdness. What worse is that AI Slop a regression in realism, and falls onto the left side of the valley.

Although we as humans can identify this slop, it is currently surprisingly hard for an AI to do so, although it hasn't stopped people from trying to build AIs that can detect AIs which in practice is filled with false positives that hurt real creatives. For slop-creators, this is a feature: if an AI company released a tool to reliably detect and punish slop, it would make their generative AI less valuable. It's reported that one of the reasons that OpenAI won't release a reliable ChatGPT text detector is that it could harm their business.

The core reason for the big tech companies allowing generative AI to cause the enshittification of the internet is misaligned incentives between the companies hosting AI slop and the users viewing it. Social media companies and their shareholders care about North Star metrics such as user retention and time-on-site, and normally those metrics can be correlated with user happiness and satisfaction with the service. But time-on-site, for example, can *also* be maximized by making the site harder and slower to use, and the deluge of AI slop accomplishes that. AI companies typically don't have analytics tracking negative user sentiment about their use of AI: if anything, the uncompromising backlash against AI convinces the companies that complainers are just a lost demographic to accommodate and double down on what they're already doing. Aggregate metrics treat human-made content and AI-generated content as equal, but *humans* do not.

Generative AI, even for researchers and practitioners such as myself, is a heavily nuanced topic that is very difficult to communicate succinctly, more difficult to do on social media which highly discourages nuance and context, and *even more difficult* as AI hypesters muddy the waters with misleading praises of generative AI such that they're easy to dunk on which just gets them more engagement and revenue. "Made by AI" is now a term that inspires dread, far from the Keaton Patti days where made-by-AI was an indicator of joyful weirdness. Bashing AI is now a meme, and there's isn't a single potential AI project that could challenge that perception because the well is poisoned beyond repair.

Would a 24/7 AI-Generated Twitch Stream Even Work Anymore?

How does the modern AI backlash tie back into AI Seinfeld? Twitch's core demographic is the same demographic as those most against the use of generative AI. Part of the reason AI Seinfeld became so successful on Twitch is because of the community it cultivated: it wouldn't have gone viral if people weren't spamming microwave `MMM`s and and answering what did the fish say when it hit the wall. Even though Twitch viewers are mostly lurkers and not chatters, a channel with a good community builds word-of-mouth even outside of Twitch, which is how Twitch channels go viral.

I decided to determine what it would take to produce a "fixed" AI Seinfeld in 2024, given both the advances in AI and the ethics involved. Now, it's definitely not anything a scrappy group of hackers could do anymore. Sure, you could once again ask an LLM to generate a sitcom script and get a bunch of assets from the Unity Asset Store, but *that's already been done before*. In order to overcome the reflexive assumption that new AI generated content is slop, the stream would have to be something completely novel and unexpected: you can't, for example, just do an AI Curb Your Enthusiasm.

The script would be unique following from my demo of detailed parametric prompts, but it would require production-studio-class tracking and documentation for how the prompts and their parameters are used to codify said uniqueness. The stream video would still need to be rendered in Unity or another engine, but in order to be unique it would require commissioning human-made visuals and sound effects: given the animosity against those who work with AI, most artists would not accept those commissions even if they were paid at a significant premium. [5] The voices would still have to be from an existing text-to-

speech voice provider: voice cloning is right out, even with explicit consent and compensation for the voice actors.

And even if all the assets were fully sourced ethically with transparent documentation for the entire pipeline, the stream's Twitch chat would likely be derailed by `AI 👏 ART 👏 IS 👏 THEFT` spam, preventing the establishment of any community, and strict moderation to curb the spam risks causing a <u>Streisand effect</u>.

The only entities that could feasibly create a 24/7 AI-generated livestream with fully ethically-sourced content would be, ironically, the big AI companies such as OpenAI which can afford to pay licenses for said data. Even <u>Disney</u>, which owns more than enough IP to train generative models of all modalities, would never do an AI Seinfeld-esque livestream for <u>brand safety</u> reasons alone: the nonzero possibility of a Disney character unexpectedly saying something problematic during the stream would make the entire project a complete nonstarter.

What's the deal with the uncanny valley?

One of the common criticisms about generative AI pointed out by creatives is "if AI is trained on all human works, then how can it create anything new"? AI Seinfeld is the perfect counterargument: even though it's powered by a LLM, the *humans* behind it are what made it go viral. Even before ChatGPT, generative AI has always excelled as a tool. The microwave gag and the 144p visual filter were not AI-generated or an attempt to emulate aspects of the Seinfeld sitcom: they were distinct creative decisions that made the entire project more interesting, and they aren't something that you could prompt an AI to suggest to add. AI Seinfeld in hindsight was an ethical form of AI-generated media: it did not replace Seinfeld the TV show, no one would stop watching streams of Seinfeld in favor of the AI-generated alternative, and copyright holders and Jerry Seinfeld did not lose revenue due to AI Seinfeld's existence: if anything, the nostalgic buzz increased streams of the original show.

With the current trajectory of AI slop and the perverse incentives by large tech companies to not address it, I am pessimistic that AI content will ever be at a state where it will cross that final hump of the uncanny valley curve into full acceptance, and even more pessimistic about the backlash against generative AI ever subsiding. With generative model training now at the point where it requires exponentially more compute and data for increasingly marginal returns, it will take years if at all for generative AI output to reach the

far right of the uncanny valley chart, and unless the large tech companies actually create an AGI, they are unlikely to obtain higher acceptability than AI Seinfeld ever did.

I wrote most of this blog post weeks ago but held off publishing it because new AI news kept happening. Most notably, the creators of Stable Diffusion just released the FLUX.1 series of generative image AI models, which presents substantially improved coherence both to the provided prompt and within the image itself. Some of the variants are open-source, allowing the community to finetune them. The XLabs-AI/flux-RealismLora in particular focuses on realism as it name implies, and one demo from that finetune went megaviral.

One of the viral realism demo images: it does not have a dreamy look as other AI images but contextually expected stage lighting, the background and lanyard text is legible despite the depth-of-field blur, and body proportions are mostly correct except the long fingers. via /u/Glittering-Football9 on Reddit / StableDiffusion subreddit.

That example in my opinion is more real than Sora but given the mixed reactions to the image, it's right at the acceptability = 0 threshold.

The generative AI bell cannot be unrung. As you can tell from this post, I personally try to thread the thin line between both cool applications of generative AI (at the risk of getting harrassed) and the problems generative AI can cause (also at the risk of getting harrassed) because it's important to shine a light on what's actually possible with AI when the misinformation around generative AI is only increasing. It's overall a big bummer how we went from weird Valentine's Day hearts, to a quirky livestream of a group of AI-generated friends, to what AI is now.

1. All of the examples in this post use LLM APIs as they provide the customization necessary to get effective results: the results for asking the same prompts to free chat frontends such as chatgpt.com will be substantially different. ↵
2. When I was younger, I actually didn't like Seinfeld and instead preferred to watch Everybody Loves Raymond. ↵
3. Incidentally, parametric prompts is why Unlimited Steam got permanently banned from Twitch: in what would now be known as a prompt injection, one of the GitHub-hosted lists the channel sourced thousands of food choices for the prompt contained a few highly offensive selections. ↵

4. Prompt engineering instability grows exponentially as the prompt size increases since each part of the prompt has to relate to each other. Claude 3.5 Sonnet is the first LLM I've tested that can handle super-long bespoke prompts and can actually account for all aspects of the prompt. ↵

5. To be fully ethical, an AI practitioner would have to proactively offer additional contractual guarantees to creatives they are commissioning, including highly-scoped usage of the assets they provide and a clause to not train generative AI on said assets to avoid future business. ↵

Max Woolf (@minimaxir) is a Data Scientist at BuzzFeed in San Francisco who works with AI/ML tools and open source projects. *Max's projects are funded by his Patreon.*

ChatGPT   Text Generation

NEXT »

The Super Effectiveness of Pokémon Embeddings Using Only Raw JSON and Images