

How generative AI could reinvent what it means to play

AI-powered NPCs that don't need a script could make games—and other worlds—deeply immersive.

By Niall Firth
June 20, 2024



GEORGE WYLESOL

Listen to this article  00:00 / 27:46

To hear more, [download the Noa app](#)

First, a confession. I only got into playing video games a little over a year ago (I know, I know). A

Christmas gift of an Xbox Series S “for the kids” dragged me—pretty easily, it turns out—into the world of late-night gaming sessions. I was immediately attracted to open-world games, in which you’re free to explore a vast simulated world and choose what challenges to accept. Red Dead Redemption 2 (RDR2), an open-world game set in the Wild West, blew my mind. I rode my horse through sleepy towns, drank in the saloon, visited a vaudeville theater, and fought off bounty hunters. One day I simply set up camp on a remote hilltop to make coffee and gaze down at the misty valley below me.

To make them feel alive, open-world games are inhabited by vast crowds of computer-controlled characters. These animated people—called NPCs, for “nonplayer characters”—populate the bars, city streets, or space ports of games. They make these virtual worlds feel lived in and full. Often—but not always—you can talk to them.



ROCKSTAR GAMES



ROCKSTAR GAMES

In open-world games like Red Dead Redemption 2, players can choose diverse interactions within the same simulated experience.

After a while, however, the repetitive chitchat (or threats) of a passing stranger forces you to bump up against the truth: *This is just a game*. It’s still fun—I had a whale of a time, honestly, looting stagecoaches, fighting in bar brawls, and stalking deer through rainy woods—but the illusion starts to weaken when you poke at it. It’s only natural. Video games are carefully crafted objects, part of a multibillion-dollar industry, that are designed to be consumed. You play them, you loot a few stagecoaches, you finish, you move on.

It may not always be like that. Just as it is upending other industries, generative AI is opening the door to entirely new kinds of in-game interactions that are open-ended, creative, and unexpected. The game may not always have to end.

Startups employing generative-AI models, like ChatGPT, are using them to create characters that don’t rely on scripts but, instead, converse with you freely. Others are experimenting with NPCs who appear to have entire interior worlds, and who can continue to play even when you, the player, are not around to watch. Eventually, generative AI could create game experiences that are infinitely detailed, twisting and changing every time you experience them.

The field is still very new, but it’s extremely hot. In 2022 the venture firm Andreessen Horowitz launched Games Fund, a \$600 million fund dedicated to gaming startups. A huge number of these are planning to use AI in gaming. And the firm, also known as A16Z, has now invested in two studios that are aiming to create their own versions of AI NPCs. A second \$600 million round was announced in April 2024.

Related Story



These six questions will dictate the future of generative AI

Generative AI took the world by storm in 2023. Its future—and ours—will be shaped by what we do next.

Early experimental demos of these experiences are already popping up, and it may not be long before they appear in full games like RDR2. But some in the industry believe this development will not just make future open-world games incredibly immersive; it could change what kinds of game worlds or experiences are even possible. Ultimately, it could change what it means to play.

“What comes after the video game? You know what I mean?” says Frank Lantz, a game designer and director of the NYU Game Center. “Maybe we’re on the threshold of a new kind of game.”

These guys just won't shut up

The way video games are made hasn't changed much over the years. Graphics are incredibly realistic. Games are bigger. But the way in which you interact with characters, and the game world around you, uses many of the same decades-old conventions.

"In mainstream games, we're still looking at variations of the formula we've had since the 1980s," says Julian Togelius, a computer science professor at New York University who has a startup called Modl.ai that does in-game testing. Part of that tried-and-tested formula is a technique called a dialogue tree, in which all of an NPC's possible responses are mapped out. Which one you get depends on which branch of the dialogue tree you have chosen. For example, say something rude about a passing NPC in RDR2 and the character will probably lash out—you have to quickly apologize to avoid a shootout (unless that's what you want).

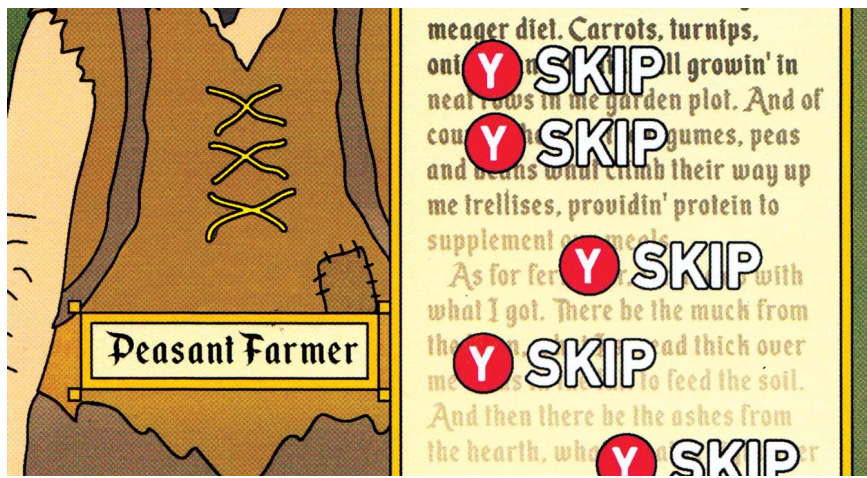
In the most expensive, high-profile games, the so-called AAA games like Elden Ring or Starfield, a deeper sense of immersion is created by using brute force to build out deep and vast dialogue trees. The biggest studios employ teams of hundreds of game developers who work for many years on a single game in which every line of dialogue is plotted and planned, and software is written so the in-game engine knows when to deploy that particular line. RDR2 reportedly contains an estimated 500,000 lines of dialogue, voiced by around 700 actors.

"You get around the fact that you can [only] do so much in the world by, like, insane amounts of writing, an insane amount of designing," says Togelius.

Generative AI is already helping take some of that drudgery out of making new games. Jonathan Lai, a general partner at A16Z and one of Games Fund's managers, says that most studios are using image-generating tools like Midjourney to enhance or streamline their work. And in a 2023 survey by A16Z, 87% of game studios said they were already using AI in their workflow in some way—and 99% planned to do so in the future. Many use AI agents to replace the human testers who look for bugs, such as places where a game might crash. In recent months, the CEO of the gaming giant EA said generative AI could be used in more than 50% of its game development processes.

Ubisoft, one of the biggest game developers, famous for AAA open-world games such as Assassin's Creed, has been using a large-language-model-based AI tool called Ghostwriter to do some of the grunt work for its developers in writing basic dialogue for its NPCs. Ghostwriter generates loads of options for background crowd chatter, which the human writer can pick from or tweak. The idea is to free the humans up so they can spend that time on more plot-focused writing.





GEORGE WYLESOL

Ultimately, though, everything is scripted. Once you spend a certain number of hours on a game, you will have seen everything there is to see, and completed every interaction. Time to buy a new one.

But for startups like Inworld AI, this situation is an opportunity. Inworld, based in California, is building tools to make in-game NPCs that respond to a player with dynamic, unscripted dialogue and actions—so they never repeat themselves. The company, now valued at \$500 million, is the best-funded AI gaming startup around thanks to backing from former Google CEO Eric Schmidt and other high-profile investors.

Role-playing games give us a unique way to experience different realities, explains Kyran Gibbs, Inworld's CEO and founder. But something has always been missing. "Basically, the characters within there are dead," he says.

"When you think about media at large, be it movies or TV or books, characters are really what drive our ability to empathize with the world," Gibbs says. "So the fact that games, which are arguably the most advanced version of storytelling that we have, are lacking these live characters—it felt to us like a pretty major issue."

Gamers themselves were pretty quick to realize that LLMs could help fill this gap. Last year, some came up with ChatGPT mods (a way to alter an existing game) for the popular role-playing game Skyrim. The mods let players interact with the game's vast cast of characters using LLM-powered free chat. One mod even included OpenAI's speech recognition software Whisper AI so that players could speak to the players with their voices, saying whatever they wanted, and have full conversations that were no longer restricted by dialogue trees.

The results gave gamers a glimpse of what might be possible but were ultimately a little disappointing. Though the conversations were open-ended, the character interactions were stilted, with delays while ChatGPT processed each request.

Inworld wants to make this type of interaction more polished. It's offering a product for AAA game studios in which developers can create the brains of an AI NPC that can be then imported into their game. Developers use the company's "Inworld Studio" to generate their NPC. For example, they can fill out a core description that sketches the character's personality, including likes and dislikes, motivations, or useful backstory. Sliders let you set levels of traits such as introversion or extroversion, insecurity or confidence. And you can also use free text to make the character drunk, aggressive, prone to exaggeration—pretty much anything.

Developers can also add descriptions of how their character speaks, including examples of commonly used phrases that Inworld's various AI models, including LLMs, then spin into dialogue in keeping with the character.

“Because there’s such reliance on a lot of labor-intensive scripting, it’s hard to get characters to handle a wide variety of ways a scenario might play out, especially as games become more and more open-ended.”

Jeff Orkin, founder, Bitpart

Game designers can also plug other information into the system: what the character knows and doesn’t know about the world (no Taylor Swift references in a medieval battle game, ideally) and any relevant safety guardrails (does your character curse or not?). Narrative controls will let the developers make sure the NPC is sticking to the story and isn’t wandering wildly off-base in its conversation. The idea is that the characters can then be imported into video-game graphics engines like Unity or Unreal Engine to add a body and features. Inworld is collaborating with the text-to-voice startup ElevenLabs to add natural-sounding voices.

Inworld’s tech hasn’t appeared in any AAA games yet, but at the Game Developers Conference (GDC) in San Francisco in March 2024, the firm unveiled an early demo with Nvidia that showcased some of what will be possible. In Covert Protocol, each player operates as a private detective who must solve a case using input from the various in-game NPCs. Also at the GDC, Inworld unveiled a demo called NEO NPC that it had worked on with Ubisoft. In NEO NPC, a player could freely interact with NPCs using voice-to-text software and use conversation to develop a deeper relationship with them.

LLMs give us the chance to make games more dynamic, says Jeff Orkin, founder of Bitpart AI, a new startup that also aims to create entire casts of LLM-powered NPCs that can be imported into games. “Because there’s such reliance on a lot of labor-intensive scripting, it’s hard to get characters to handle a wide variety of ways a scenario might play out, especially as games become more and more open-ended,” he says.

Bitpart’s approach is in part inspired by Orkin’s PhD research at MIT’s Media Lab. There, he trained AIs to role-play social situations using game-play logs of humans doing the same things with each other in multiplayer games.

Bitpart’s casts of characters are trained using a large language model and then fine-tuned in a way that means the in-game interactions are not entirely open-ended and infinite. Instead, the company uses an LLM and other tools to generate a script covering a range of possible interactions, and then a human game designer will select some. Orkin describes the process as authoring the Lego bricks of the interaction. An in-game algorithm searches out specific bricks to string them together at the appropriate time.

Bitpart’s approach could create some delightful in-game moments. In a restaurant, for example, you might ask a waiter for something, but the bartender might overhear and join in. Bitpart’s AI currently works with Roblox. Orkin says the company is now running trials with AAA game studios, although he won’t yet say which ones.

But generative AI might do more than just enhance the immersiveness of existing kinds of games. It could give rise to completely new ways to play.

Making the impossible possible

When I asked Frank Lantz about how AI could change gaming, he talked for 26 minutes straight. His initial reaction to generative AI had been visceral: “I was like, oh my God, this is my destiny and is what I was put on the planet for.”

Lantz has been in and around the cutting edge of the game industry and AI for decades but

received a cult level of acclaim a few years ago when he created the Universal Paperclips game. The simple in-browser game gives the player the job of producing as many paper clips as possible. It's a riff on the famous thought experiment by the philosopher Nick Bostrom, which imagines an AI that is given the same task and optimizes against humanity's interest by turning all the matter in the known universe into paper clips.

Lantz is bursting with ideas for ways to use generative AI. One is to experience a new work of art as it is being created, with the player participating in its creation. "You're inside of something like *Lord of the Rings* as it's being written. You're inside a piece of literature that is unfolding around you in real time," he says. He also imagines strategy games where the players and the AI work together to reinvent what kind of game it is and what the rules are, so it is never the same twice.

For Orkin, LLM-powered NPCs can make games unpredictable—and that's exciting. "It introduces a lot of open questions, like what you do when a character answers you but that sends a story in a direction that nobody planned for," he says.

Generative AI might do more than just enhance the immersiveness of existing kinds of games. It could give rise to completely new ways to play.

It might mean games that are unlike anything we've seen thus far. Gaming experiences that unspool as the characters' relationships shift and change, as friendships start and end, could unlock entirely new narrative experiences that are less about action and more about conversation and personalities.

Togelius imagines new worlds built to react to the player's own wants and needs, populated with NPCs that the player must teach or influence as the game progresses. Imagine interacting with characters whose opinions can change, whom you could persuade or motivate to act in a certain way—say, to go to battle with you. "A thoroughly generative game could be really, really good," he says. "But you really have to change your whole expectation of what a game is."

Lantz is currently working on a prototype of a game in which the premise is that you—the player—wake up dead, and the afterlife you are in is a low-rent, cheap version of a synthetic world. The game plays out like a noir in which you must explore a city full of thousands of NPCs powered by a version of ChatGPT, whom you must interact with to work out how you ended up there.

His early experiments gave him some eerie moments when he felt that the characters seemed to know more than they should, a sensation recognizable to people who have played with LLMs before. Even though you know they're not alive, they can still freak you out a bit.

"If you run electricity through a frog's corpse, the frog will move," he says. "And if you run \$10 million worth of computation through the internet ... it moves like a frog, you know."

But these early forays into generative-AI gaming have given him a real sense of excitement for what's next: "I felt like, okay, this is a thread. There really is a new kind of artwork here."

If an AI NPC talks and no one is around to listen, is there a sound?

AI NPCs won't just enhance player interactions—they might interact with one another in weird ways. *Red Dead Redemption 2*'s NPCs each have long, detailed scripts that spell out exactly where they should go, what work they must complete, and how they'd react if anything unexpected occurred. If you want, you can follow an NPC and watch it go about its day. It's fun, but ultimately it's hard-coded.

NPCs built with generative AI could have a lot more leeway—even interacting with one another

when the player isn't there to watch. Just as people have been fooled into thinking LLMs are sentient, watching a city of generated NPCs might feel like peering over the top of a toy box that has somehow magically come alive.

We're already getting a sense of what this might look like. At Stanford University, Joon Sung Park has been experimenting with AI-generated characters and watching to see how their behavior changes and gains complexity as they encounter one another.

Because large language models have sucked up the internet and social media, they actually contain a lot of detail about how we behave and interact, he says.



Gamers came up with ChatGPT mods for the popular role-playing game Skyrim.

COURTESY OF THE DEVELOPERS



Although 2016's hugely hyped No Man's Sky used procedural generation to create endless planets to explore, many saw it as a letdown.

COURTESY OF THE DEVELOPERS



In Covert Protocol, players operate as private detectives who must solve the case using input from various in-game NPCs

COURTESY OF THE DEVELOPERS

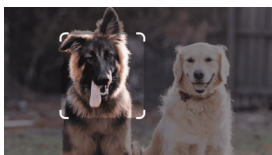
In Park's recent research, he and colleagues set up a Sims-like game, called Smallville, with 25 simulated characters that had been trained using generative AI. Each was given a name and a simple biography before being set in motion. When left to interact with each other for two days, they began to exhibit humanlike conversations and behavior, including remembering each other and being able to talk about their past interactions.

For example, the researchers prompted one character to organize a Valentine's Day party—and then let the simulation run. That character sent invitations around town, while other members of the community asked each other on dates to go to the party, and all turned up at the venue at the correct time. All of this was carried out through conversations, and past interactions between characters were stored in their "memories" as natural language.

For Park, the implications for gaming are huge. "This is exactly the sort of tech that the gaming community for their NPCs have been waiting for," he says.

His research has inspired games like AI Town, an open-source interactive experience on GitHub that lets human players interact with AI NPCs in a simple top-down game. You can leave the NPCs to get along for a few days and check in on them, reading the transcripts of the interactions they had while you were away. Anyone is free to take AI Town's code to build new NPC experiences through AI.

Related Story



Google is throwing generative AI at everything

But experts say that releasing these models into the wild before fixing their flaws could prove extremely risky for the company.

For Daniel De Freitas, cofounder of the startup Character AI, which lets users generate and interact with their own LLM-powered characters, the generative-AI revolution will allow new types of games to emerge—ones in which the NPCs don't even need human players.

The player is "joining an adventure that is always happening, that the AIs are playing," he imagines. "It's the equivalent of joining a theme park full of actors, but unlike the actors, they truly 'believe' that they are in those roles."

If you're getting *Westworld* vibes right about now, you're not alone. There are plenty of stories about people torturing or killing their simple Sims characters in the game for fun.

Would mistreating NPCs that pass for real humans cross some sort of new ethical boundary?

What if, Lantz asks, an AI NPC that appeared conscious begged for its life when you simulated torturing it?

It raises complex questions he adds. “One is: What are the ethical dimensions of pretend violence? And the other is: At what point do AIs become moral agents to which harm can be done?”

There are other potential issues too. An immersive world that feels real, and never ends, could be dangerously addictive. Some users of AI chatbots have already reported losing hours and even days in conversation with their creations. Are there dangers that the same parasocial relationships could emerge with AI NPCs?

“We may need to worry about people forming unhealthy relationships with game characters at some point,” says Togelius. Until now, players have been able to differentiate pretty easily between game play and real life. But AI NPCs might change that, he says: “If at some point what we now call ‘video games’ morph into some all-encompassing virtual reality, we will probably need to worry about the effect of NPCs being too good, in some sense.”

A portrait of the artist as a young bot

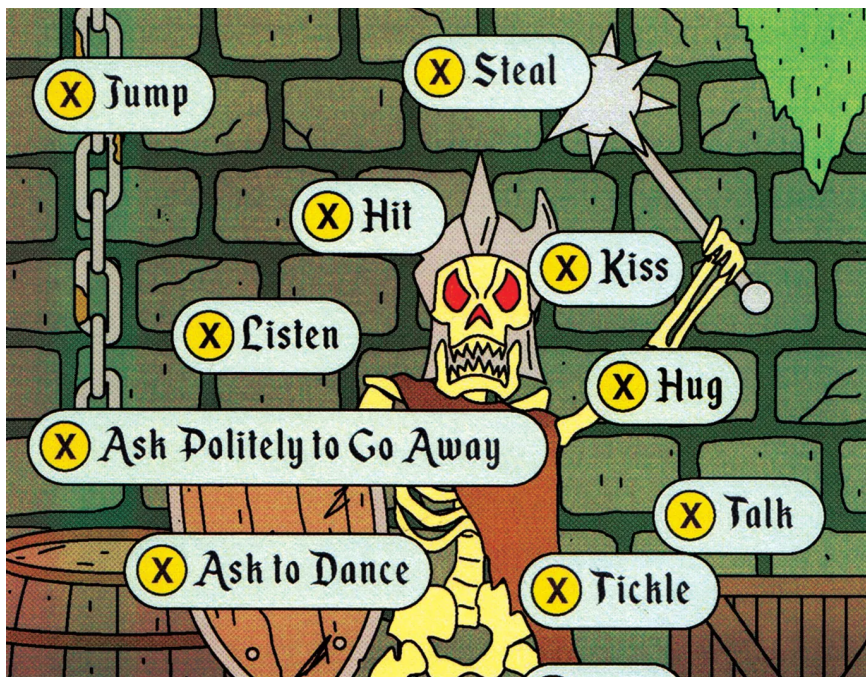
Not everyone is convinced that never-ending open-ended conversations between the player and NPCs are what we really want for the future of games.

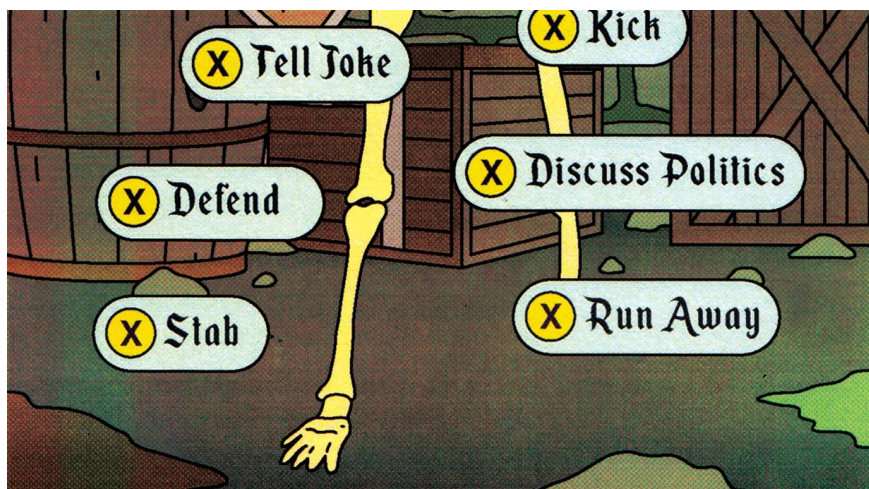
“I think we have to be cautious about connecting our imaginations with reality,” says Mike Cook, an AI researcher and game designer. “The idea of a game where you can go anywhere, talk to anyone, and do anything has always been a dream of a certain kind of player. But in practice, this freedom is often at odds with what we want from a story.”

In other words, having to generate a lot of the dialogue yourself might actually get kind of ... well, boring. “If you can’t think of interesting or dramatic things to say, or are simply too tired or bored to do it, then you’re going to basically be reading your own very bad creative fiction,” says Cook.

Orkin likewise doesn’t think conversations that could go anywhere are actually what most gamers want. “I want to play a game that a bunch of very talented, creative people have really thought through and created an engaging story and world,” he says.

This idea of authorship is an important part of game play, agrees Togelius. “You can generate as much as you want,” he says. “But that doesn’t guarantee that anything is interesting and worth keeping. In fact, the more content you generate, the more boring it might be.”





GEORGE WYLESOL

Sometimes, the possibility of everything is too much to cope with. No Man's Sky, a hugely hyped space game launched in 2016 that used algorithms to generate endless planets to explore, was seen by many players as a bit of a letdown when it finally arrived. Players quickly discovered that being able to explore a universe that never ended, with worlds that were endlessly different, actually fell a little flat. (A series of updates over subsequent years has made No Man's Sky a little more structured, and it's now generally well thought of.)

One approach might be to keep AI gaming experiences tight and focused.

Hilary Mason, CEO at the gaming startup Hidden Door, likes to joke that her work is “artisanal AI.” She *is* from Brooklyn, after all, says her colleague Chris Foster, the firm's game director, laughing.

Hidden Door, which has not yet released any products, is making role-playing text adventures based on classic stories that the user can steer. It's like Dungeons & Dragons for the generative AI era. It stitches together classic tropes for certain adventure worlds, and an annotated database of thousands of words and phrases, and then uses a variety of machine-learning tools, including LLMs, to make each story unique. Players walk through a semi-unstructured storytelling experience, free-typing into text boxes to control their character.

The result feels a bit like hand-annotating an AI-generated novel with Post-it notes.

In a demo with Mason, I got to watch as her character infiltrated a hospital and attempted to hack into the server. Each suggestion prompted the system to spin up the next part of the story, with the large language model creating new descriptions and in-game objects on the fly.

Each experience lasts between 20 and 40 minutes, and for Foster, it creates an “expressive canvas” that people can play with. The fixed length and the added human touch—Mason's artisanal approach—give players “something really new and magical,” he says.

There's more to life than games

Park thinks generative AI that makes NPCs feel alive in games will have other, more fundamental implications further down the line.

“This can, I think, also change the meaning of what games are,” he says.

For example, he's excited about using generative-AI agents to simulate how real people act. He thinks AI agents could one day be used as proxies for real people to, for example, test out the likely reaction to a new economic policy. Counterfactual scenarios could be plugged in that would let policymakers run time backwards to try to see what would have happened if a different path had been taken.

“You want to learn that if you implement this social policy or

Related Story**An early guide to policymaking on generative AI**

How lawmakers are thinking about the risks of the latest tech revolution

economic policy, what is going to be the impact that it's going to have on the target population?" he suggests. "Will there be unexpected side effects that we're not going to be able to foresee on day one?"

And while Inworld is focused on adding immersion to video games, it has also worked with LG in South Korea to make characters that kids can chat with to improve their English language skills. Others are using Inworld's tech to create interactive experiences. One of these, called Moment in Manzanar, was created to help players empathize with the Japanese-Americans the US government detained in internment camps during World War II. It allows the user to

speak to a fictional character called Ichiro who talks about what it was like to be held in the Manzanar camp in California.

Inworld's NPC ambitions might be exciting for gamers (my future excursions as a cowboy could be even more immersive!), but there are some who believe using AI to enhance existing games is thinking too small. Instead, we should be leaning into the weirdness of LLMs to create entirely new kinds of experiences that were never possible before, says Togelius. The shortcomings of LLMs "are not bugs—they're features," he says.

Lantz agrees. "You have to start with the reality of what these things are and what they do—this kind of latent space of possibilities that you're surfing and exploring," he says. "These engines already have that kind of a psychedelic quality to them. There's something trippy about them. Unlocking that is the thing that I'm interested in."

Whatever is next, we probably haven't even imagined it yet, Lantz thinks.

"And maybe it's not about a simulated world with pretend characters in it at all," he says.

"Maybe it's something totally different. I don't know. But I'm excited to find out."