

The Generative AI Revolution in Games

by James Gwertzman and Jack Soslow

To understand how radically gaming is about to be transformed by Generative AI, look no further than this recent [Twitter post](#) by [@emmanuel_2m](#). In this post he explores using Stable Diffusion + Dreambooth, popular 2D Generative AI models, to generate images of potions for a hypothetical game.



What's transformative about this work is not just that it saves time and money while also delivering quality – thus smashing the classic “you can only have two of cost, quality, or speed” triangle. Artists are now creating high-quality images in a matter of hours that would otherwise take weeks to generate by hand. What's truly transformative is that:

- This creative power is now available to anyone who can learn a few simple tools.
- These tools can create an endless number of variations in a highly iterative way.
- Once trained, the process is real-time – results are available near instantaneously.

There hasn't been a technology this revolutionary for gaming since real-time 3D. Spend any time at all talking to game creators, and the sense of excitement and wonder is palpable. So where is this technology going? And how will it transform gaming? First, though, let's review what is Generative AI?

What is Generative AI

Generative AI is a category of machine learning where computers can generate original new content in response to prompts from the user. Today text and images are the most mature applications of this technology, but there is work underway in virtually every creative domain, from animation, to sound effects, to music, to even creating virtual characters with fully fleshed out personalities.

AI is nothing new in games, of course. Even early games, like Atari's Pong, had computer-controlled opponents to challenge the player. These virtual foes, however, were not running AI as we know it today. They were simply scripted procedures crafted by game designers. They simulated an artificially intelligent opponent, but they couldn't learn, and they were only as good as the programmers who built them.

What's different now is the amount of computing power available, thanks to faster microprocessors and the cloud. With this power, it's possible to build large neural networks that can identify patterns and representations in highly complex domains.

This blog post has two parts:

- Part I consists of our observations and predictions for the field of Generative AI for games.
- Part II is our market map of the space, outlining the various segments and identifying key companies in each.

Part I – Observations and Predictions

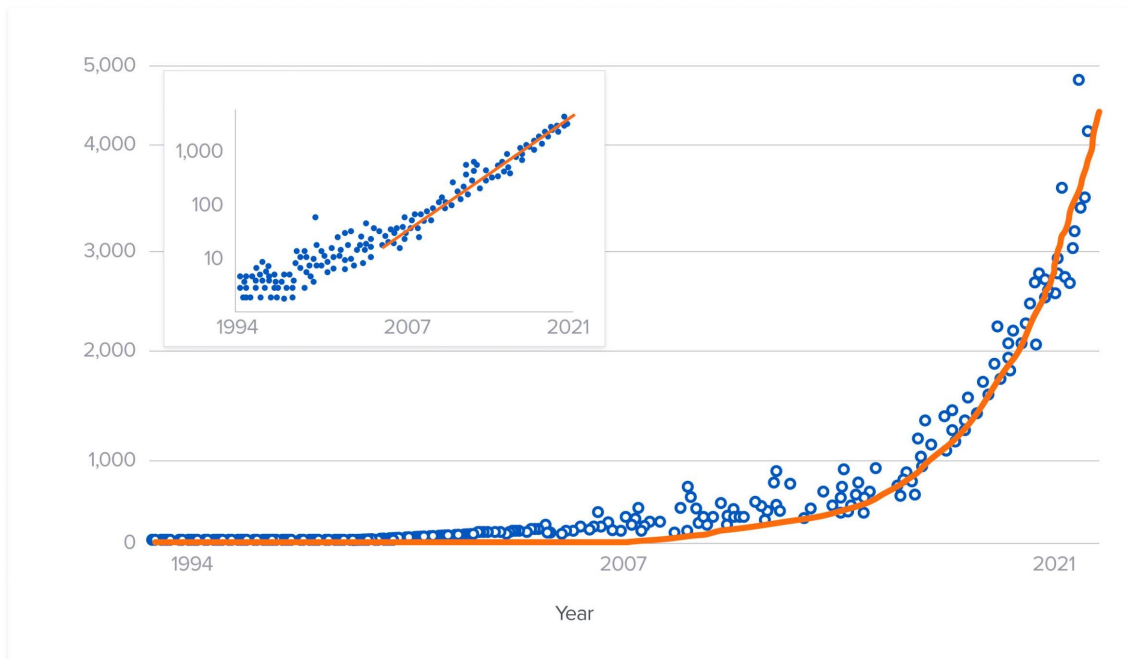
Assumptions

First, let's explore some assumptions underlying the rest of this blog post:

1. The amount of research being done in general AI will continue to grow, creating ever more effective techniques

Consider this graph of the number of academic papers published on Machine Learning or Artificial Intelligence in the [arXiv archive](#) each month:

ML + AI arXiv Papers Per Month



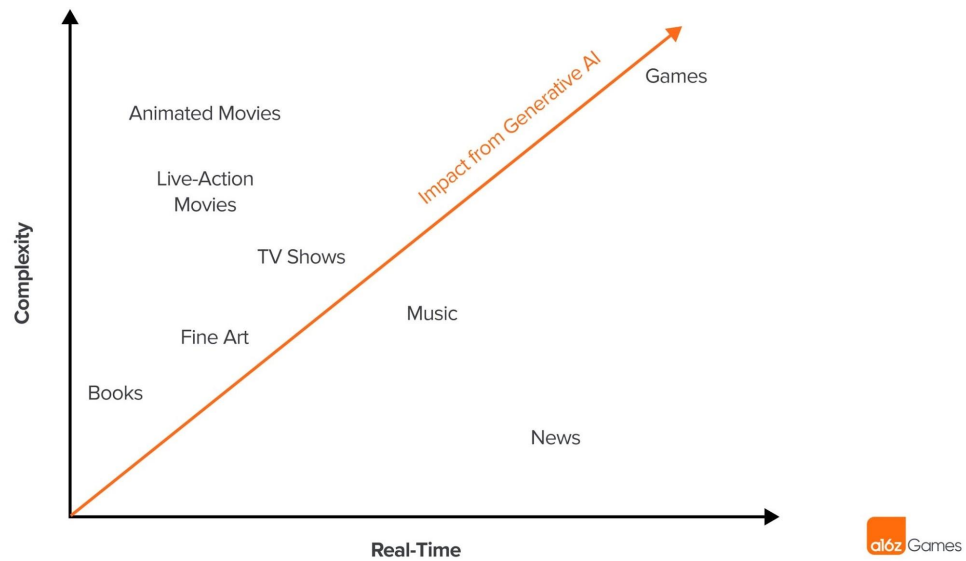
Source: arXiv



As you can see, the number of papers is growing exponentially, with no sign of slowing down. And this just includes published papers – much of the research is never even published, going directly to open source models or product R&D. The result is an explosion in interest and innovation.

2. Of all entertainment, games will be most impacted by Generative AI

Games are the most complex form of entertainment, in terms of the sheer number of asset types involved (2D art, 3D art, sound effects, music, dialog, etc). Games are also the most interactive, with a heavy emphasis on real-time experiences. This creates a steep barrier to entry for new game developers, as well as a steep cost to produce a modern, chart-topping game. It also creates a tremendous opportunity for Generative AI disruption.



Consider a game like Red Dead Redemption 2, one of the most expensive games ever produced, costing nearly \$500 million to make. It's easy to see why – it has one of the most beautiful, fully realized virtual worlds of any game on the market. It also took nearly 8 years to build, features more than 1,000 non-playable characters (each with its own personality, artwork, and voice actor), a world nearly 30 square miles in size, more than 100 missions split across 6 chapters, and almost 60 hours of music created by over 100 musicians. Everything about this game is big.



Now compare Red Dead Redemption 2 to Microsoft Flight Simulator, which is not just big, it's enormous. Microsoft Flight Simulator enables players to fly around the

entire planet Earth, all 197 million square miles of it. How did Microsoft build such a massive game? By letting an AI do it. Microsoft partnered with blackshark.ai, and trained an AI to generate a photorealistic 3D world from 2D satellite images.



This is an example of a game that would have literally been impossible to build without the use of AI, and furthermore, benefits from the fact that these models can be continually improved over time. For example, they can enhance the “highway cloverleaf overpass” model, re-run the entire build process, and suddenly all the highway overpasses on the entire planet are improved.

3. There will be a generative AI model for every asset involved in game production

So far 2D image generators like Stable Diffusion, or MidJourney have captured the majority of the popular excitement over Generative AI due to the eye-catching nature of the images they can generate. But already there are Generative AI models for virtually all assets involved in games, from 3D models, to character animations, to dialog and music. The second half of this blog post includes a market map highlighting some of the companies focusing on each type of content.

4. The price of content will drop dramatically, going effectively to zero in some cases.

When talking to game developers who are experimenting with integrating Generative AI into their production pipeline, the greatest excitement is over the dramatic

reduction in time and cost. One developer has told us that their time to generate concept art for a single image, start to finish, has dropped down from 3 weeks to a single hour: a 120-to-1 reduction. We believe similar savings will be possible across the entire production pipeline.

To be clear, artists are not in danger of being replaced. It does mean that artists no longer need to do all the work themselves: they can now set initial creative direction, then hand off much of the time consuming and technical execution to an AI. In this, they are like cel painters from the early days of hand-drawn animation in which highly skilled “inkers” drew the outlines of animation, and then armies of lower-cost “painters” would do the time-consuming work of painting the animation cels, filling in the lines. It’s the “auto-complete” for game creation.

5. We are still in the infancy of this revolution and a lot of practices still need to be refined

Despite all the recent excitement, we are still just at the starting line. There is an enormous amount of work ahead as we figure out how to harness this new technology for games, and enormous opportunities will be generated for companies who move quickly into this new space.

Predictions

Given these assumptions, here are some predictions for how the game industry may be transformed:

1. Learning how to use Generative AI effectively will become a marketable skill

Already we are seeing some experimenters using Generative AI more effectively than others. To make the most use of this new technology requires using a variety of tools and techniques and knowing how to bounce between them. We predict this will become a marketable skill, combining the creative vision of an artist with the technical skills of a programmer.

Chris Anderson is famous for saying, “Every abundance creates a new scarcity.” As content becomes abundant, we believe it’s the artists who know how to work most collaboratively and effectively with the AI tools who will be in the most short supply.

For example, to use Generative AI for production artwork carries special challenges, including:

- **Coherence.** With any production asset, you need to be able to make changes or edits to the asset down the road. With an AI tool, that means needing to be able to reproduce the asset with the same prompt, so you can then make changes. This can be tricky as the same prompt can generate vastly different outcomes.
- **Style.** It's important for all art in a given game to have a consistent style – which means your tools need to be trained on or otherwise tied to your given style.

2. Lowering barriers will result in more risk-taking and creative exploration

We may soon be entering a new “golden age” of game development, in which a lower barrier to entry results in an explosion of more innovative and creative games. Not just because lower production costs result in lower risk, but because these tools unlock the ability to create high-quality content for broader audiences. Which leads to the next prediction...

3. A rise in AI-assisted “micro game studios”

Armed with Generative AI tools and services, we will start to see more viable commercial games produced by tiny “micro studios” of just 1 or 2 employees. The idea of a small indie game studio is not new – hit game Among Us was created by studio Innersloth with just 5 employees – but the size and scale of the games these small studios can create will grow. This will result in...

4. An Increase in the number of games released each year

The success of Unity and Roblox have shown that providing powerful creative tools result in more games being built. Generative AI will lower the bar even further, creating an even greater number of games. The industry already suffers from discovery challenges – more than 10,000 games were added to Steam last year alone – and this will put even more pressure on discovery. However we will also see...

5. New game types created that were not possible before Generative AI

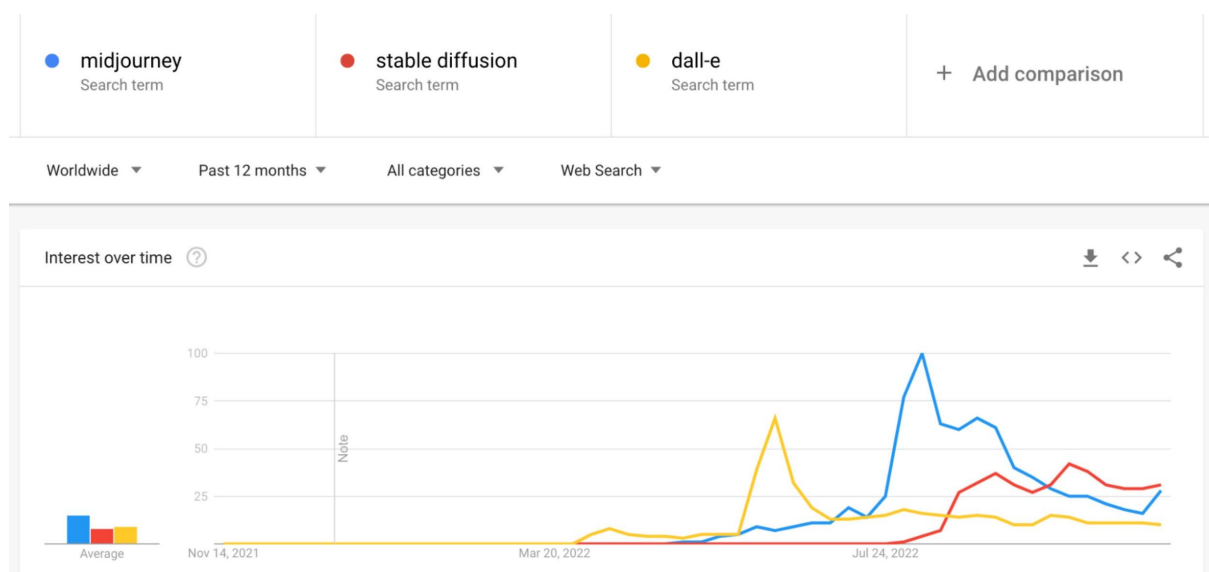
We will be seeing new game genres invented that were simply not possible without Generative AI. We already talked about Microsoft's flight simulator, but there will be entirely new genres invented that depend on real-time generation of new content.

Consider Arrowmancer, by Spellbrush. This is an RPG game that features AI-created characters for virtually unlimited new gameplay.

We also know of another game developer that is using AI to let players create their own in-game avatar. Previously they had a collection of hand-drawn avatar images that players could mix-and-match to create their avatar – now they have thrown this out entirely, and are simply generating the avatar image from the player’s description. Letting players generate content through an AI is safer than letting players upload their own content from scratch, since the AI can be trained to avoid creating offensive content, while still giving players a greater sense of ownership.

6. Value will accrue to industry specific AI tools, and not just foundational models

The excitement and buzz around foundational models like Stable Diffusion and Midjourney are generating eye-popping valuations, but the continuing flood of new research ensures that new models will come and go as new techniques are refined. Consider website search traffic to 3 popular Generative AI models: Dall-E, Midjourney, and Stable Diffusion. Each new model has its turn in the spotlight.



An alternative approach may be to build industry aligned suites of tools that focus on the Generative AI needs of a given industry, with deep understanding of a particular audience, and rich integration into existing production pipelines (such as Unity or Unreal for games).

A good example is Runway which targets the needs of video creators with AI assisted tools like video editing, green screen removal, inpainting, and motion

tracking. Tools like this can build and monetize a given audience, adding new models over time. We have not yet seen a suite such as Runway for games emerge yet, but we know it's a space of active development.

7. Legal challenges are coming

What all of these Generative AI models have in common is that they are trained using massive datasets of content, often created by scraping the Internet itself. Stable Diffusion, for example, is trained on more than 5 billion image/caption pairs, scraped from the web.

At the moment these models are claiming to operate under the “fair use” copyright doctrine, but this argument has not yet been definitively tested in court. It seems clear that legal challenges are coming which will likely shift the landscape of Generative AI.

It's possible that large studios will seek competitive advantage by building proprietary models built on internal content they have clear right & title to. Microsoft, for example, is especially well positioned here with 23 first party studios today, and another 7 after its acquisition of Activision closes.

8. Programming will not be disrupted as deeply as artistic content – at least not yet

Software engineering is the other major cost of game development, but as our colleagues on the a16z Enterprise team have shared in their recent blog post, Art Isn't Dead. It's Just Machine-Generated, generating code with an AI model requires more testing and verification, and thus has a smaller productivity improvement than generating creative assets. Coding tools like Copilot may provide moderate performance improvements for engineers, but won't have the same impact... at least anytime soon.

Recommendations

Based on these predictions, we offer the following recommendations:

1. Start exploring Generative AI now

It's going to take a while to figure out how to fully leverage the power of this coming Generative AI revolution. Companies that start now will have an advantage later. We know several studios who have internal experimental projects underway to explore how these techniques can impact production.

2. Look for market map opportunities

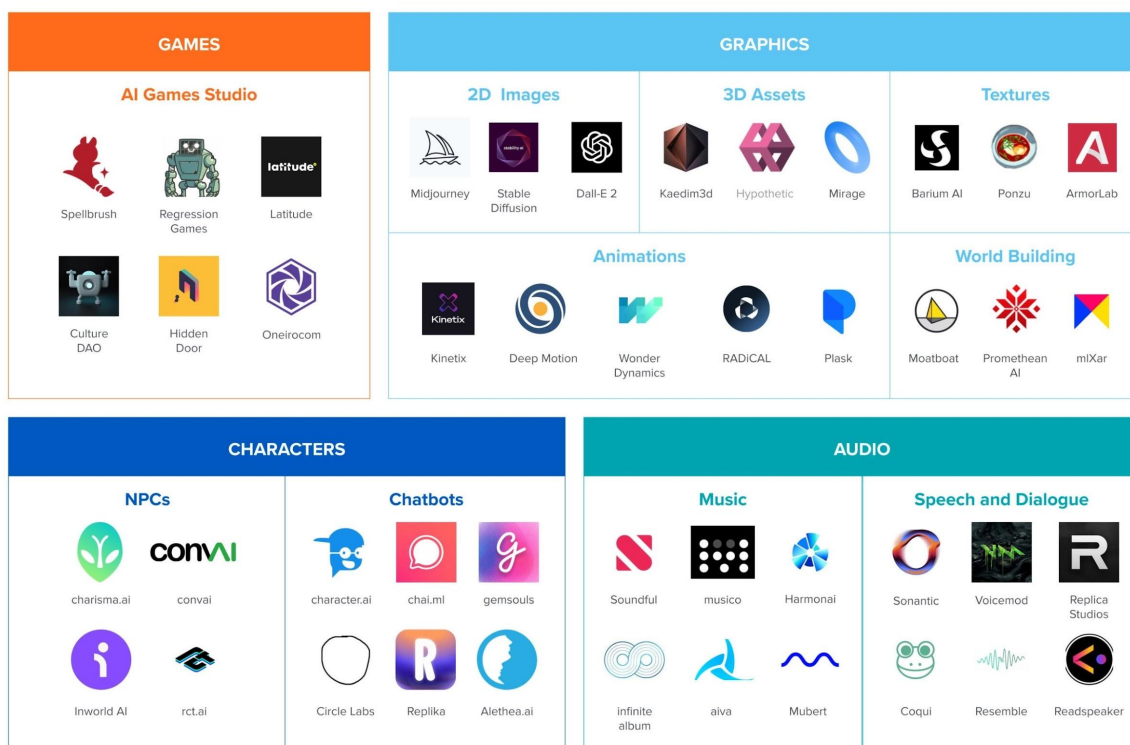
Some parts of our market map are very crowded already, like Animations or Speech & Dialog, but other areas are wide open. We encourage entrepreneurs interested in this space to focus their efforts on the areas that are still unexplored, such as “Runway for Games”.

Part II – Market Map

Current state of the market

We have created a market map to capture a list of the companies we’ve identified in each of these categories where we see Generative AI impacting games. This blog post goes through each of those categories, explaining it in a bit more detail, and highlighting the most exciting companies in each category.

Generative AI for Games Market Map



Graphics provided herein are for informational purposes only and should not be relied upon when making any investment decision.



2D Images

Generating 2D images from text prompts is already one of the most widely applied areas of generative AI. Tools like Midjourney, Stable Diffusion, and Dall-E 2 can

generate high quality 2D images from text, and have already found their way into game production at multiple stages of the game life cycle.

Concept Art

Generative AI tools are excellent at “ideation” or helping non-artists, like game designers, explore concepts and ideas very quickly to generate concept artwork, a key part of the production process. For example, one studio (staying anonymous) is using several of these tools together to radically speed up their concept art process, taking a single day to create an image that previously would have taken as long as 3 weeks.

- First, their game designers use Midjourney to explore different ideas and generate images they find inspiring.
- These get turned over to a professional concept artist who assembles them together and paints over the result to create a single coherent image – which is then fed into Stable Diffusion to create a bunch of variations.
- They discuss these variations, pick one, paint in some edits manually – then repeat the process until they’re happy with the result.
- At that stage, then pass this image back into Stable Diffusion one last time to “upscale” it to create the final piece of art.

2D Production Art

Some studios are already experimenting with using the same tools for in-game production artwork. For example, here is a nice [tutorial from Albert Bozesan](#) on using Stable Diffusion to create in-game 2D assets.

3D Artwork

3D assets are the building block of all modern games, as well as the upcoming metaverse. A virtual world, or game level, is essentially just a collection of 3D assets, placed and modified to populate the environment. Creating a 3D asset, however, is more complex than creating a 2D image, and involves multiple steps including creating a 3D model and adding textures and effects. For animated characters, it also involves creating an internal “skeleton”, and then creating animations on top of that skeleton.

We’re seeing several different startups going after each stage of this 3D asset creation process, including model creation, character animation, and level building.

This is not yet a solved problem, however – none of the solutions are ready to be fully integrated into production yet.

3D assets

Startups trying to solve the 3D model creation problem include [Kaedim](#), [Mirage](#), and [Hypothetic](#). Larger companies are also looking at the problem, including Nvidia's [Get3D](#) and Autodesk's [ClipForge](#). Kaedim and Get3d are focused on image-to-3D; ClipForge and Mirage are focused on text-to-3D, while Hypothetic is interested in both text-to-3D search, as well as image-to-3D.

3D Textures

A 3D model only looks as realistic as the texture or materials that are applied to the mesh. Deciding which mossy, weathered stone texture to apply to a medieval castle model can completely change the look and feel of a scene. Textures contain metadata on how light reacts to the material (i.e. roughness, shininess, etc). Allowing artists to easily generate textures based on text or image prompts will be hugely valuable towards increasing iteration speed within the creative process. Several teams are pursuing this opportunity including [BariumAI](#), [Ponzu](#), and [ArmorLab](#).

Animation

Creating great animation is one of the most time consuming, expensive, and skillful parts of the game creation process. One way to reduce the cost, and to create more realistic animation, is to use motion capture, in which you put an actor or dancer in a motion capture suit and record them moving in a specially instrumented motion capture stage.

We're now seeing Generative AI models that can capture animation straight from a video. This is much more efficient, both because it removes the need for an expensive motion capture rig, and because it means you can capture animation from existing videos. Another exciting aspect of these models is that they can also be used to apply filters to existing animations, such as making them look drunk, or old, or happy. Companies going after this space include [Kinetix](#), [DeepMotion](#), [RADiCAL](#), [Move Ai](#), and [Plask](#).

Level design & world building

One of the most time consuming aspects of game creation is building out the world of a game, a task that generative AI should be well suited to. Games like Minecraft,

No Man's Sky, and Diablo are already famous for using procedural techniques to generate their levels, in which levels are created randomly, different every time, but following rules laid down by the level designer. A big selling point of the new Unreal 5 game engine is its collection of procedural tools for open world design, such as foliage placement.

We've seen a few initiatives in the space, like [Promethean](#), [MLXAR](#), or Meta's [Builder Bot](#), and think it's only a matter of time before generative techniques largely replace procedural techniques. There has been academic research in the space for a while, including [generative techniques for Minecraft](#) or [level design in Doom](#).

Another compelling reason to look forward to generative AI tools for level design would be the ability to create levels and worlds in different styles. You could imagine asking tools to generate a world in 1920's flapper era New York, vs dystopian blade-runner-esque future, vs. Tolkien-esque fantasy world.

The following concepts were generated by Midjourney using the prompt, "a game level in the style of..."



Audio

Sound and music are a huge part of the gameplay experience. We're starting to see companies using Generative AI to generate audio to complement the work already happening on the graphics side.

Sound Effects

Sound effects are an attractive open area for AI. There have been [academic papers](#) exploring the idea of using AI to generate “foley” in film (e.g. footsteps) but few commercial products in gaming yet.

We think this is only a matter of time, since the interactive nature of games make this an obvious application for generative AI, both creating static sound effects as part of production (“laser gun sound, in the style of Star Wars”), and creating real-time interactive sound effects at run-time.

Consider something as simple as generating footstep sounds for the player’s character. Most games solve this by including a small number of pre-recorded footstep sounds: walking on grass, walking on gravel, running on grass, running on gravel, etc. These are tedious to generate and manage, and sound repetitive and unrealistic at runtime.

A better approach would be a real-time generative AI model for foley sound effects, that can generate appropriate sound effects, on the fly, slightly differently each time, that are responsive to in-game parameters such as ground surface, weight of character, gait, footwear, etc.

Music

Music has always been a challenge for games. It’s important, since it can help set the emotional tone just as it does in film or television, but since games can last for hundreds or even thousands of hours, it can quickly become repetitive or annoying. Also, due to the interactive nature of games, it can be hard for the music to precisely match what’s happening on screen at any given time.

Adaptive music has been a topic in game audio for more than two decades, going all the way back to Microsoft’s “[DirectMusic](#)” system for creating interactive music. DirectMusic was never widely adapted, due largely to the difficulty of composing in the format. Only a few games, like Monolith’s [No One Lives Forever](#), created truly interactive scores.

Now we’re seeing a number of companies trying to create AI generated music, such as [Soundful](#), [Musico](#), [Harmonai](#), [Infinite Album](#), and [Aiva](#). And while some tools today, like [Jukebox](#) by Open AI, are highly computationally intensive and can’t run in real-time, the majority can run in real-time once the initial model is built.

Speech and Dialog

There are a large number of companies trying to create realistic voices for in-game characters. This is not surprising given the long history of trying to give computers a voice through speech synthesis. Companies include [Sonantic](#), [Coqui](#), [Replica Studios](#), [Resemble.ai](#), [Readspeaker.ai](#), and many more.

There are multiple advantages to using generative AI for speech, which partly explains why this space is so crowded.

- **Generate dialog on-the-fly.** Typically speech in games is pre-recorded from voice actors, but these are limited to pre-recorded canned speeches. With generative AI dialog, characters can say anything – which means they can fully react to what players are doing. Combined with more intelligent AI models for NPC's (outside the scope of this blog, but an equally exciting area of innovation right now), the promise of games that are fully reactive to players are coming soon.
- **Role playing.** Many players want to play as fantasy characters that bear little resemblance to their real-world identity. This fantasy breaks down, however, as soon as players speak in their own voices. Using a generated voice that matches the player's avatar maintains that illusion. Control. As the speech is generated, you can control the nuance of the voice like its timbre, inflection, emotional resonance, phoneme length, accents, and more.
- **Localization.** Allows dialog to be translated into any language and spoken in the same voice. Companies like [Deepdub](#) are focused specifically on this niche.

NPCs or player characters

Many startups are looking at using generative AI to create believable characters you can interact with, partly because this is a market with such wide applicability outside of games, such as virtual assistants or receptionists.

Efforts to create believable characters go back to the beginnings of AI research. In fact, the definition of the classic “Turing Test” for artificial intelligence is that a human should be unable to distinguish between a chat conversation with an AI versus a human.

At this point there are hundreds of companies building general purpose chatbots, many of them powered by the GPT-3 like language models. A smaller number are specifically trying to build chatbots for the purpose of entertainment, such as [Replika](#) and [Anima](#) who are trying to build virtual friends. The concept of dating a virtual girlfriend, as explored in the movie *Her*, may be closer than you think.

We are now seeing the next iteration of these chatbot platforms, such as [Charisma.ai](#), [Convai.com](#), or [Inworld.ai](#), meant to power fully rendered 3D characters, with emotions, and agency, with tools to allow the creator to give these characters goals. This is important if they're going to fit within a game or have a narrative place in advancing the plot forward, versus purely being window dressing.

All-in-one platforms

One of the most successful generative AI tools at large is [Runwayml.com](#), because it brings together a broad suite of creator tools in a single package. Currently there is no such platform serving video games, and we think this is an overlooked opportunity. We would love to invest in a solution that features:

- Full set of generative AI tools covering the entire production process. (code, asset generation, textures, audio, descriptions, etc.)
- Tightly integrated with popular game engines like Unreal and Unity.
- Designed to fit into a typical game production pipeline.

Conclusion

This is an incredible time to be a game creator! Thanks in part to the tools described in this blog post, it has never been easier to generate the content needed to build a game – even if your game is as large as the entire planet!

It's even possible to one day imagine an entire personalized game, created just for the player, based on exactly what the player wants. This has been in science fiction for a long time – like the “AI Mind Game” in Ender’s Game, or the holodeck in Star Trek. But with the tools described in this blog post advancing as quickly as they are, it's not hard to imagine this reality is just around the corner.

If you are a founder, or potential founder, interested in building an AI for Gaming company, please reach out! We want to hear from you!