

Google for Startups

Future of AI

Perspectives on
generative media
for startups.



Table of contents

<u>Foreword</u>	<u>03</u>
<u>Predictions about generative media</u>	<u>04</u>
<u>Perspectives of industry leaders</u>	<u>07</u>
↳ Victor Riparbelli Co-Founder and CEO, Synthesia	↳ Darian Shirazi Managing Partner, Gradient
↳ Joaquín Cuenca Abela Co-Founder and CEO, Freepik	↳ David Benjamin Team Lead, Creative Lab and Co-Founder, AI Futures Fund, Google
↳ Grace Isford Partner, Lux Capital	↳ Grace Wang Co-Founder and CMO, OpusClip
↳ Jaclyn Konzelmann Director of Product Management, Google Labs	↳ Sami Ede Co-Founder and Senior ML Researcher, Leonardo.Ai
<u>Build with Google's generative media models</u>	<u>32</u>



The 2026 opportunity

Innovation has always been defined by the bet. That is, the courage to pursue an idea before the rest of the world sees its value. And in today's generative era, the window to prove value has narrowed significantly. We are witnessing an unprecedented compression of the "idea-to-scale" cycle. It is no longer enough to simply build; founders must now navigate a landscape where AI is the core, not just a wrapper.

As we look toward the future of generative media and creative disciplines, the winners will be those who move beyond basic orchestration to create deep, specialized moats. To thrive, we must embed AI into the fabric of our workflows and focus on solving high-value, real-world problems. We are moving toward a world of "vibe design," where you don't need to be a coder or a professional editor to bring a vision to life. If you can describe it, if you can feel it, you can build it. This shift democratizes creativity, allowing a solo entrepreneur to run a photoshoot for a global brand from their laptop, or a student to transform 60 research sources into a deep-dive audio conversation.

This report provides a range of perspectives of the future possibilities of AI in creative disciplines, straight from founders that are redefining the space, investors that are betting on them, and Google experts building the foundational tools. As you read through these findings, you will see that the next wave of innovation belongs to those who prioritize scalability and differentiation from day one. I hope these perspectives inspire you to begin building what's next.



Darren Mowry

Vice President, Global Startups, Google



Predictions about generative media

Video will replace text.

As the marginal cost of creating video content drops to zero, enterprises will embrace it over traditional formats like slides and docs. Why? Because humans prefer to consume visual content.



Victor Riparbelli
Co-Founder and CEO,
Synthesia

The film industry will flip to AI.

Within the next two to three years, AI-driven media production will overtake traditional methods economically, collecting more revenue than traditional films as the industry flips to this new way of creating.



Joaquín Cuenca Abela
Co-Founder and CEO,
Freepik

We're moving toward a post-keyboard world.

Traditional interfaces and form factors will disappear in favor of neural and brain-computer interfaces that intuit thoughts and extend the human mind.



Grace Isford
Partner,
Lux Capital

Founders will become creative directors.

But instead of stepping back, this shift will make them the most hands-on builders in history—multiplying their output by directing a fleet of agents to get things done.



Jaclyn Konzelmann
Director of Product Management,
Google Labs

Individual creators will dominate the media landscape.

Content creators will become more relevant than traditional media brands. These creators can research topics deeply, and deliver content to consumers in meaningful, personalized ways.



Darian Shirazi

Managing Partner,
Gradient

Gamers will enter 'living' metaverses.

Real-time generative AI will transform gaming into a responsive, 'living' metaverse that adapts its narrative and atmosphere to a player's unique interests.



David Benjamin

Team Lead, Creative Lab and
Co-Founder, AI Futures Fund, Google

Personalized videos will lead the way.

The biggest shift for AI startups will be moving from 'novelty' to 'personalization.' The next generation of tools won't just generate generic video, they will precisely understand an individual's style and enhance their specific story.



Grace Wang

Co-Founder and CMO,
OpusClip

AI will soon generate haptics and spatial acoustics.

AI is extremely good at learning, and it won't be long before it's creating spatial acoustics for 3D experiences. And, as tailored actuators become more readily available, it also won't be long before AI-generated haptics can also become part of a brand's toolkit.



Sami Ede

Co-Founder and Senior ML Researcher,
Leonardo.Ai



Perspectives of industry leaders

As AI-made videos become ubiquitous, work is transformed

A new genre in video has arrived, and it's opening up a whole new world for builders.

Are we witnessing the end of static content? As builders continue to push the boundaries of what's possible with AI in video content creation, the idea of static content will take a backseat to just-in-time content generated just for you.

Similar to how EDM shook up the music industry, we're on the cusp of a revolution in video. And, just like drum machines and synthesizers didn't replace traditional music genres and real instruments, the arrival of AI-generated videos won't replace traditional media. Instead, it creates opportunity.

For startups wanting to play a part in this revolution, the key lies in thinking about what consumers want and need from content—whether they are at home on the couch or in the office at work.



Victor Riparbelli

Co-Founder and CEO, Synthesia

Synthesia is the world's leading AI video platform for business, with visual communication and enterprise skill development products to help people work better.

Originally from Denmark, Victor co-founded the London-based company in 2017 after seeing the potential of computer vision to democratize media creation. Recognized on the 2024 TIME100 AI list, he is a prominent voice in the global shift toward using generative media as a primary medium for business communication, skill development, and knowledge sharing.



How agentic video will play out in the enterprise

There's no doubt that AI is transforming how we create content. With the latest models, anyone with the tools can create virtually anything. The question is, how is this impacting work?

More A/V, less ABC

We'll consume more audio and video content than text. If you think about office jobs today, most of the communication is still in slides or written documents—mainly because it's cheaper and easier to produce content in these formats.

In the next few years, the cost gap between creating a text document and video in the workplace will disappear. We'll enter a world where 45-second videos replace long slide decks, making training programs a lot easier to digest. And where video and audio will replace a lot of the content on a B2B website, enhancing the browsing experience.

Personalization plus

As history proves, every time we create new media technologies, we create new media formats that are native to those technologies.

So now, we're moving into a world where every video can be truly personalized to the person watching it. We can generate the video first and foremost, but can also put an LLM on top of it to make a decision on how to personalize the video specific to that person. And we don't need a human in the loop to do this.

Given all this, I think the experience of media in the future is going to be very different to what it is today—much more multimodal and rendered in real time for whoever's watching it.

Avatars for everyone

Humans are attracted to faces. Being able to look at another person and take in their expression is a really important part of communication.

Avatars will step into a new role here. They have evolved so much compared to early iterations, and will soon form part of a person's digital identity. You'll have an avatar version of yourself; an LLM of yourself. And you'll use it in all sorts of communication.

Already, we're seeing people use avatars so they don't have to get in front of the camera all the time. For the enterprise, they'll prove useful in other ways, too. For example, say you've created a big library of video content for your customers, but the 'face' of the videos leaves the company. If it's an AI video, you just switch out the avatar to someone else.

“

The idea of static content will take a backseat to just-in-time content generated just for you.”



AI colleagues

With AI systems being tuned in real-time, video agents and AI colleagues will come to the fore.

For example, think of a sales training situation where a video agent takes the employee through a role play experience with a customer prospect. Or an AI colleague who talks you through your employee feedback survey.

We're increasingly moving towards a world where these AI colleagues become part of the lifecycle of the company.



You'll have an avatar version of yourself; an LLM of yourself. And you'll use it in all sorts of communication."

Advice for founders

When building, test ideas with utility

With everything changing so fast, startups need to get into experimentation mode. What are the right formats? How will people want to engage with this stuff?

When you're experimenting, prioritize utility over novelty. Sure, you could build the cool, interesting thing with a high 'wow' factor that everyone talks about. But if it doesn't have a strong real-world application, then it just remains a fun toy. You want it to become a real product.

Remember, some things will take off; others won't. Enjoy the creative process.

Making movies takes on a whole new meaning with AI

The ability to tell a great story is going to be the defining feature of filmmaking and content creation in the AI era.

It won't be long before the craft of filmmaking will be distilled down into a single element: storytelling.

With AI doing much of the heavy lifting in terms of production and execution, films will soon be made by individuals—similar to how a single person writes a book. While the process of creating the film may take time, just like writing a book does, it's not going to be a technical challenge. Rather, filmmakers can focus on the story; and on shaping that story into something that really resonates with people.

Here, I want to share some of the questions that are top of my mind—and those of the people I'm talking to day-to-day—as we navigate these new waters.



**Joaquín
Cuenca Abela**

Co-Founder and CEO, Freepik

Freepik is an AI-powered creative platform serving over 100 million users worldwide. Under Joaquín's leadership, it has scaled from a three-person startup into a global leader integrating millions of visual assets and advanced generative tools for image, video, and audio creation.

A serial entrepreneur, Joaquín previously co-founded Panoramio, which became one of the first Spanish startups to be acquired by Google in 2007.



Will films as we know them disappear?

The short answer: no. Higher-end films will still be produced, with around 50 people working on them. These people will shape the story and craft the characters, drawing on unique human insight to bring them to life.

It's the execution side of filmmaking that is more likely to disappear and be replaced by AI. And, as part of this, we're going to see many more use cases for videos that we can't even conceive of today.

We're already seeing phenomenal expansion of AI video creation. It's been used all over the place—and I think that, this year, we will see more productions made by AI versus traditional filmmaking. In the next two or three years, it will take over economically, with AI productions collecting more revenue than traditional productions.

There are a few verticals, like animation, where we simply can't keep doing it the old way. So, this year, animation will completely flip to a new model.

Does AI have taste?

Another short (and controversial) answer: yes. With AI doing more and more, there's a lot of talk about 'taste' being this high-gear element that AI can't touch. I disagree. AI will have taste—and we will be able to tilt this taste into different directions. Indeed, AI is already showing us its creative side in the images and footage it's producing today.

But while AI has taste, you can easily copy it. You can't do that with a human. You can't plug something in and get a perfect copy of someone else's mind. Rather, humans have this unique, uncopyable element—our past experiences and everything that shaped us into who we are.

With this in mind, my view is that, in the future, there will be economic value and there will be value in taking our experiences and pouring them into what we create.

And that's where we differ from AI. Whereas AI creates content based on an echo of the past, humans create things by drawing on our unique, lived experiences. We must pour these experiences into our work to make something that you don't find in a machine. Doing so will help us create things that others hopefully find novel and interesting.

“

There's a lot of talk about 'taste' being this high-gear element that AI can't touch. I disagree. AI will have taste—and we will be able to tilt this taste into different directions.”



How should we think about copyright in this new era?

As we move forward, the industry must become more pragmatic about how AI tools are trained. Until now, people have focused on whether the AI companies have permission to use publicly available data to train their models. We're rapidly moving into a position where it's acceptable to use anything under fair use.

With this in mind, we need to shift the conversation from copyright on the input images to responsibility for the output. That is, we need to make sure that whatever is produced by the AI is legally usable. Remember, AI is a tool. Like any tool, it can be used to violate IP by accident, so our role is to help users identify and avoid those situations.



Travel, meet new people, learn new things—you'll gather a ton of ideas that will become more valuable than any algorithm. Think of it like AI—it gives you, personally, a bigger training data set.”

Advice for founders

How do you build a winning startup?

Based on my observations above and all the other things I'm seeing in the industry today, startups need to focus on differentiating characteristics. My advice is to:

01

Go out in the wild

To build a meaningful startup, explore diverse lived experiences. Travel, meet new people, learn new things—you'll gather a ton of ideas that will become more valuable than any algorithm. Think of it like AI—it gives you, personally, a bigger training data set.

02

Think laterally

The best ideas come from being a little bit out of the mainstream. It might be that you take a proven idea and use it in a different context.

03

Keep an open mind

My biggest mistake in the past was believing that creativity was unique to humans. It isn't. Many things can be replicated mechanically. But the specific way you combine your life, your struggles, and your perspective into a message—that is something you won't find in a machine.

Reimagining how we interact with technology

What once felt futuristic—like AI becoming a second brain—is almost here and now.

What will the future of user and consumer experiences look like? How will we interact with AI? As the world moves further into the AI era, questions like these are top of mind for investors like myself.

With things moving so fast, we need builders who have real intentionality about how we interact with AI. We need builders who dream up totally new interfaces that can intuit and interpret our needs; those who design completely new modalities and mechanisms for AI-driven user experiences.

It's a period of extraordinary change in the media industry, and founders are ideally poised to seize upon the new opportunities that this evolving landscape presents.



Grace Isford

Partner, Lux Capital

As a Partner in Lux's New York City office, Grace Isford invests in companies at the nexus of data, AI infrastructure, AI applications, and AI in the physical world—from pre-seed to growth.

Before joining Lux, Grace was a principal at Canvas Ventures, where she started as a campus scout while at Stanford University before becoming a full-time investor in 2019. Prior to Canvas, Grace worked on the LP side at the Stanford Management Company, in product at Handshake, and in growth equity at Stripes Group.



People still oversimplify AI, thinking of it in terms of chat bots or image bots. Very few understand how different our lives will be when generative AI is used in entirely new modalities.

If we reimagine the future of human-computer interaction, we can see a new world forming—one in which AI and your mind are married together in a really seamless way. For example, we're starting to see some very interesting frontier advancements in the brain-computer interface space. It's a whole new modality, capable of intuiting and interpreting neural signals to then become an extension of your mind. In this future, we won't need to articulate what we want done—the neural device will just know.

While all this can feel far off, I believe we're getting closer. And we'll soon be at the point where we're not using our keyboards anymore, and just using our voice to get things done. We're on the cusp of completely new interfaces, personalized to every user.

New modalities and the future of compute

These new interfaces will have significantly higher compute needs. Even the video models that we're seeing today are a lot more compute intensive and require a different set of tools. We'll need new video infrastructure companies to serve this demand.

Likewise, new interfaces will impact how we approach AI models. I believe technical architectures will become more focused on 'mini' models or smaller models that can do more with less, and can work in swarms. This approach is a lot more efficient from a compute perspective, plus there's the benefit of being able to run smaller models on a device. I'm excited to see the impact of these smaller models ultimately democratizing access to AI, making it cheaper and easier for more people to use the product.

“

We're starting to see some very interesting frontier advancements in the brain-computer interface space. It's a whole new modality, capable of intuiting and interpreting neural signals to then become an extension of your mind.”



How future-focused founders can grab investors' attention

As we forge this new future for media, my advice to founders is to focus on the following key traits.

01

Be opinionated about your product

Create something that feels authentic and unique to you. I particularly admire when founders have a point of view and a vision of how the world should be—and how the tools they are building fit into this world. I may not agree with their point of view, but I'm always looking for founders' unique insights, their earned insights, and how they think about the market in a way that other folks may not.

02

Know what good taste means for your users

Taste is the moat that matters. The best builders understand what good taste looks like in the eye of their users. Where this gets complicated is that taste is subjective, and so 'good' for one app and its users may not be so good for another. It all comes back to knowing your users really well.

03

Persistently challenge the status quo

The best founders are very inquisitive, constantly challenging the status quo and asking why things are the way they are. With this in mind, never assume that a given interface or a given product will continue to exist.

04

Know your technology

In my experience, the best builders have a deep understanding of both tech stacks and creative requirements. This is particularly important right now, because we're still in an early phase of the tech growth of the sector—with AI only just starting to penetrate many of these creative companies.

05

Know your niche

Find the go-to-market angle that's unique to you and most relevant to your product. Don't try to create a product for everyone that does everything—sure, it could be great, but it's going to be really hard to find users and get quick feedback. Instead, start with a hypothesis, then test and iterate until you have a more precise vision to scale with.

06

Ultimately, build a great product

My best advice is to double down on creating value for your customers by building a great product that they really need, care about, and will pay for. Whether you're in a bull market, a construction market, a bear market or beyond, this is the key to success.

“

I particularly admire when founders have a point of view, an opinion, and a vision of how the world should be—and how the tools they are building fit into this world.”

Building AI tools with a generative media lens

The latest generative media models are opening up vast new possibilities for AI builders.

Despite all the breakthrough moments we've had with AI so far, we're only at the beginning of the journey. Every new model shines a light on how much further we can go with it all.

Without being able to predict the future, one thing is clear: AI is fundamentally changing the way things work. We're already doing things that were never possible before. And, because of this, we need to rethink a lot of what we've built so far and figure out if these things should continue to exist in the world or if they should take a different shape or form.



Jaclyn Konzelmann

Director of Product Management,
Google Labs

Jaclyn leads teams building the next generation of AI-first products. Specializing in the 0-to-1 stage, she launched Google AI Studio and the Gemini API, making Google's most capable models widely accessible. She also led Project Mariner, exploring human-agent interaction.

Currently, her portfolio spans diverse products—from Pomelli, which helps SMBs generate on-brand content, to experiments like Opal and Mixboard that reimagine creating with generative AI.



Purpose-built AI applications are not only changing how people express themselves, but how we get things done.

A generative media lens is now being applied to how we build products and experiences for non-creative use cases.

For example, in education, students no longer need to follow a one-size-fits-all, linear path in a textbook. Instead, they can access personalized lessons delivered across different media and modalities to suit their learning style and tastes. The ability to lean into generative media and create these immersive experiences on an individual basis is really exciting to think about.

Another example is world models, which carry so much potential for helping people master new skills through simulations. Already, people can learn how to drive a car and fly a plane, but now we're seeing opportunities for all types of learning to be accessed in simulated worlds.

Ultimately, AI will enable us to go back to spending quality time with other people, doing things that may not have been possible before. When you think about all the crazy, wild things you want to do in life, they're probably all possible—but just take a lot of time or money. AI accelerates what's possible, reducing the barrier for creating real-world experiences. The question is, how can we apply all this to a better world?



As good as AI can be at executing things, humans still have an amazing role to play in providing the taste, the vision, and the decision on what should be built or created. Just because anything can be built, it doesn't mean everything should be."

AI will have a huge impact on creativity

It's exciting to think about how generative media can apply to things like education and everyday tasks. But what about its impact on creativity?

Creative tools are going to look and feel a lot different in the future. Previously, the limiting factor would be that people didn't have the right tools, or they didn't know how to use the tools.

These traditional barriers to entry are rapidly dissolving. If I want to design a cute basket for my kid's bike, I can. If I want to create a watercolor painting, but don't know how to paint, I can. This ability for everyone to start doing things with new AI tools is going to unlock the next wave of creativity.

Then there are those that don't have the time for creativity, like busy solopreneurs or two-person startups. For these small businesses, generative media tools will transform how they create on-brand marketing assets. We've built a tool, Pomelli, that's geared towards exactly this.

Pomelli

Pomelli is an AI marketing tool that helps you generate scalable, tailored, on-brand content to grow your business. The marketing creatives generated with Pomelli can be downloaded and used across all social media platforms.

[Learn more](#)



Humans will play a key role in directing AI models

It never ceases to surprise me that, even when my AI agent presents me with a good idea, I still want to take it ten steps further—iterating and refining the work until I am completely happy with it. In this way, humans are becoming creative directors. It's on us to figure out what should be created; it's also on us to guide the AI and work with it to bring our vision to life.

As humans become more familiar with AI tools and their possibilities, our creative vision will widen. We'll start exploring more comprehensive, complex, and imaginative things. And we'll rely on agents to do the toilsome, time-consuming work that previously hampered our creativity.



As a founder, how do you take a step back from the hands-on work that you're used to, and allow the AI to become the worker instead? How do you step into more of a creative director role?"

Advice for founders

How founders can tap into the director mindset

As a founder, how do you take a step back from the hands-on work that you're used to, and allow the AI to become the worker instead? How do you step into more of a creative director role?

Start by rolling up your sleeves and building things. Choose projects that personally resonate. For example, I always have around 12 side projects on the go at any one time, and they all help me push the limits of what's possible.

When you're ready to start applying your learnings in a business context, identify a goal. Work with the AI and try different things until you land on a repeatable pattern. Then, you can codify it, test it, and iterate some more. Over time, you'll end up with a robust solution.

If you feel stuck at any time, talk to Gemini or your AI. By articulating your roadblocks and ideating on the problem, the AI can help get you unstuck.

Lastly, keep persevering. Don't assume that because you hit a dead end, you should stop. Your agent could help you find a new path forward. Granted, by the end of the process, you may have a workflow that looks very different to what you'd planned at the beginning. That's OK. If it works, you could be onto a winner.

In the AI era, media is becoming more customized and individually relevant

AI models are unlocking new forms of creativity and new types of content, ultimately disrupting how we interact with and consume media.

The way we interact with media is changing fast. Once upon a time, we consumed information by watching TV, reading the newspaper, or using the internet. Then social media arrived, and we started getting our news there instead—while adding our own opinions to the mix.

Now, we're entering an era of bi-directional media. We communicate with a model, and it consumes information from us and the world around us to provide very specific, personalized output—all while remembering what we're interested in.

This bi-directional interaction is the future of media. We're going to consume information about the world in a way that's directly related to our world, who we are, what we care about, and what's influencing us. And it's going to evolve the media industry in new and interesting ways.



Darian Shirazi

Managing Partner, Gradient

Darian has spent his entire career leading seed investments in category-defining AI companies and working with AI-first founders.

A self-taught engineer, he founded Radius, a B2B Customer Data Platform that raised over \$100MM and counted dozens of Fortune 500 companies as customers, running it as CEO for seven years. Before that, he joined Facebook as one of its first engineers, reporting directly to Mark Zuckerberg. He studied CS and Philosophy at UC Berkeley before dropping out.



New AI models allow people to create images, videos, and animations that they could only dream of before.

Someone who may not have the motor skills to be a painter can now be a painter—if they have a vision for what that painting should be. Indeed, the more tools we can give humans to create what they imagine, the more likely we are to see original content.

Similar to how computers unlocked new forms of art and expression like memes, the future is going to reveal even more sophisticated capabilities that we haven't even thought of yet. And so, founders should be thinking, "How can I unlock new forms of human creativity?"

Whether the answer is a point solution or a more specific model for one type of creativity remains to be seen. But either way, I believe we're going to find the Picasso of this era as a result of creating new tools.

New tools will drive regional demand

These new types of creative tools will drive significant demand for local models in certain geographies; as well as the ability to serve and spin up models as needed—including on your own device.

This compute complexity places huge demand on infrastructure. We're going to see significantly more need for high-power chips and high-power data centers regionally.

Think about how much traffic social media has driven, and how these platforms created a modern economy around data centers. I think we're going to see 10x or more of that in the AI era—where consumers come from all parts of the world and are willing to pay to create their own apps, videos, images, and text. We're already starting to see it, and it's just the beginning.



We're entering an era of bi-directional media—one where we communicate with a model, and it consumes information from us and the world around us to provide very specific, personalized output."

Independent journalists will become the source of truth

In this new world, where everyone can create quality content, legacy media brands will become less relevant. Instead of going to a portal to consume your media, it will be delivered to you through the model based on your personal preferences and the individuals you think are creating the best content.

We won't trust the traditional media brands, whose left-leaning or right-leaning ideologies are often hidden by the portal we're consuming media through. Instead, we'll trust individuals that have researched topics deeply and create their own content on that topic—and independent journalism will likely thrive.

Models will help minimize misinformation

At the same time, models will get better at deciphering disinformation or misinformation from real information. Users will be able to click through for evidence, and have the model explain why it thinks something is true or false. Models naturally learn with more information and extensively cite sources of such learning—this is a major leap over social media which provides little to no evidence for even the most viral posts.

Over time, we'll have more of a grounded truth in certain aspects of content that gets published to the internet.



Bi-directional media in action

We're working with one company that has built a voice assistant model that allows anyone to communicate a memoir. You can give it to your grandparents, and they can tell their life story through this two-way script. It then combines their stories with other life stories and photos from your family, as well as other events that happened in the world.

The long-term vision is that you can communicate with your relatives after they have passed away. You'll have a digital twin of all the meaningful people in your life.

It's a form of media that we haven't had before. Anyone can write a life history in a book, but it can be hard for others to consume that information. Talking to an ancestor is a much more interesting way to not only learn their life story, but also learn about the world during that period.



Models will get better at deciphering disinformation or misinformation from real information. Users will be able to click through for evidence, and have the model explain why it thinks something is true or false.”

Advice for founders

How founders can thrive in this new media landscape

Be curious

We always seek out polymathematical, autodidactic founders—people who are always learning new things and are good at multiple things. It's about your ability to come up with something brand new, unique, and defensible.

Let go of 'good enough'

Founders should not be afraid to burn down something that is working, but not working well enough. As models get better, we're seeing businesses and ideas get disrupted more and more.

Build your moat

One way you can future-proof your startup in this rapidly evolving space is by looking at your data moat. Do you have a data moat? Do you have data network effects? Can you build a product where every new user adds value to everybody?

Tie KPIs to AI usage

Measure how your company uses AI. Do your employees need training? Could they prompt more effectively? Is data fragmentation a bottleneck? Armed with answers, you can iteratively improve AI adoption—which ultimately helps your employees do their work more effectively.

From prompt to play, AI is reshaping the gaming industry

Generative media not only levels the playing field for content creators in the gaming industry. It creates whole new playing fields.

Generative AI is the greatest creation tool in the history of humanity. Its potential is endless—helping people create things they could never make before, at significantly lower cost.

Watching it evolve is fascinating. While we're still in the silly phase of content creation with generative tools, the content is getting richer every month. Besides, you need the silly stuff—and all the silly apps—to get to the interesting stuff.

This is particularly true in gaming. Generative AI makes it easier for game creators to make content they couldn't make before, and really bring their great storylines and interesting characters to life.



David Benjamin

Team Lead, Creative Lab and
Co-Founder, AI Futures Fund, Google

David is a Team Lead in Google's Creative Lab, which helps invent Google's future and communicate the company's innovations to the world.

He is also Co-Founder and Co-Lead at Google's AI Futures Fund, which launched in 2025 to help startups at various growth stages get early access to Google DeepMind's latest AI models as well as resources, technical expertise, and potential equity funding to accelerate their progress.

At Google, David focuses on researching emerging tech, incubating experimental prototypes, and investing in/partnering with startups. He's worked on projects involving many of Google's core products, including Gemini, YouTube, Google Maps, Google Play, and Ads.



As generative media democratizes the ability to create compelling gaming experiences, new opportunities emerge.

While the gaming industry is the most valuable part of the entertainment ecosystem in terms of net revenue, it's always been hard to create great games. Creators needed proficiency in many different tools and skill sets.

Now, anyone can make a game—and not just any game, but 3D world games and, soon, even AAA games.

These massive AAA games used to cost tens, if not hundreds, of millions of dollars to produce. Not anymore. With AI handling a lot of the heavy lifting of asset creation, many elements of high-fidelity game production will be democratized. It will allow small teams and, eventually, individual creators to achieve the scale and polish previously reserved for massive gaming studios.

These creators will be on the lookout for creation tools that can effectively translate their specific vision into a finished product with speed and precision—which points to a big opportunity for startups.



Many elements of high-fidelity game production will be democratized. It will allow small teams and, eventually, individual creators to achieve the scale and polish previously reserved for massive gaming studios.”

Beyond AAA games, generative media will also help open up new genres and styles of game play:

- **Generative UGC:** User-generated content will become generative as we move beyond simple models to AI tools that allow players to architect entire worlds and consumption experiences in real time.
- **Deep personalization:** We will see the rise of ‘liquid games’ that evolve their mechanics based on player behavior, effectively rethinking traditional genre boundaries at both the indie and AAA levels.
- **Interactive 3D environments:** The tech will eventually bridge the gap between video and play, allowing users to step inside a 2D frame and explore a fully interactive 3D environment.
- **New social creation hubs:** A new wave of mini-game platforms will emerge, centered on AI-driven creation and social mechanics. On these platforms, the act of creating together is the primary loop.

In a world where every solo developer has access to AAA-level asset generation, gaming startups need to differentiate by developing human-driven, world-class narratives; and by demonstrating creative taste and recognizable IP.



Other thoughts on the future of generative media

Forward-thinking UX

The ultimate user experience is not a chat bot. With a chat bot, we still have to think about what to ask, and we get stuck with the blinking cursor problem of, “What do I type here?”

Rather, it’s a personalized, dynamic UI that morphs into the specific need you have at a given time. We’re seeing seeds of it sprout right now through vibe coding tools that add different layers over a prompt box. But to get to the point where the system knows exactly the interface you need at the right time, for the right purpose—which is essentially the holy grail of generative AI UX—we’ll need more time.



Gaming startups need to differentiate by developing human-driven, world-class narratives; and by demonstrating creative taste and recognizable IP.”

The value of creativity

We will always gravitate to great human artists. In a world where there are more and more synthetic artists, the value of human art and creativity will rise. The best human creators will continue to be the best—and will command a premium.

For those using generative media, the value will lie in rethinking what it means to make content—and thinking about new ways to create media content across the entire stack with generative AI tools.

Advice for founders

Strive for vertical excellence

Instead of building a broad, generic layer, focus on a loyal community and develop a specialized AI solution that solves their specific, high-value problems. Once you’ve nailed that, expand from there.

Creating authentic moments on social media in the AI era

With adoption of automated content creation tools growing fast, authenticity and human experience will become invaluable.

Right now, we are witnessing an explosion of video content on social media, with the creation of AI-generated videos becoming as easy as applying a camera filter effect. While these technically polished production capabilities are becoming accessible to everyone, the content often feels soulless, generic, and homogeneous.

As automation accelerates and video production becomes more affordable, 'human' elements like story judgment, taste, and unique points of view are becoming exponentially more valuable.



Grace Wang

Co-Founder and CMO, OpusClip

OpusClip is an AI-powered video editing platform that helps creators and businesses turn long-form videos into viral short clips, in one click.

Before launching the startup in 2022, Grace and her fellow co-founders led a 500-person social media talent agency, an experience that directly informed their understanding of the challenges creators face. Under her leadership, the company has grown to serve over 16 million users and secured significant strategic funding from investors like SoftBank Vision Fund 2.



As content creators strive for more authenticity, the future creative landscape will shift away from traditional skills.

With more elements of video production being automated, the editors, producers, and animators of tomorrow will look very different to what they look like today. Instead of possessing advanced editing skills, for example, an editor will instead act more like a creative director or narrative supervisor, and will get AI models to do the heavy lifting.

It means that the AI tools and models themselves will need to precisely understand their users—and know exactly how to personalize content and enhance storytelling for every creator in the business.

Opportunities for AI on social platforms

Character creation

AI will have an impact on the characters we use to create content. We'll soon see more AI actors and characters that may have huge creative teams behind them. As long as the creators themselves bring their unique personality and valuable experience, the AI actors' identities can be sustained.

Viral moments

AI needs to evolve beyond 'past knowledge' to really grasp virality. Because it is trained on existing patterns, it struggles to generate the 'net new' novelty that top creators aim for. It's possible to engineer virality to some extent by using models to research trends, discover popular opinions, and create attention-grabbing quotes. However, to create a true viral moment, you need something people have never seen before. I see an opportunity for AI in the future here.

Voiceovers

Voice cloning and translations are a really interesting space for AI. Today, we're seeing that local audiences still prefer authentic voices in films and clips, opting to read translated captions instead of listening to generated voiceovers. Why? Because AI can't quite capture the emotion and personality of original voiceovers yet. It points to a real opportunity for startups to capture market share in this space in the future.

Long-form content

Today's social platforms monetize the audience's attention, with engagement metrics like watch rate and time dominating algorithms. Right now, short-form AI videos are popular and perform really well. But I still believe there is a place for long-form content, and an opportunity for AI-made videos to make it into a 3-minute, 10-minute, or 30-minute story.

“

As automation accelerates and video production becomes more affordable, 'human' elements like story judgment, taste, and unique points of view are becoming exponentially more valuable.”



Three pillars of winning content

In this new ecosystem, three things will truly differentiate content and determine who ‘wins’ on social media.

01

Authentic storytelling

This is driven by a creator’s distinct personality and unique lived experiences. It’s about creators sharing their point of view with their audience.

02

Emotional depth

Content must resonate, creating a sense of rage or empathy; otherwise, it’s just AI slop that people won’t watch for more than a few seconds.

03

Taste

The ability to spot what’s working or what feels off is a critical gap that AI cannot yet bridge on its own.

“

With more elements of video production being automated, the editors, producers, and animators of tomorrow will look very different to what they look like today.”

Advice for founders

For your startup to succeed, hire AI builders

In my opinion, the best way for startups to scale is to hire AI builders, who essentially become the managers of AI agents.

So instead of hiring people defined by traditional functions, you might see a marketer or salesperson become an engineer. We already see this in a lot of startups where small teams can achieve huge impact. You don’t have to scale the marketing team size because the marketing engineers have built up a workflow that previously required a team of ten to accomplish—whereas now one or two marketers can do the entire workflow.

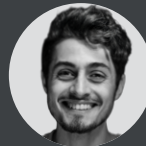
AI's transformative role in A/B testing, acoustics, 3D experiences, and more

Flat AI visuals set the foundation—now richer modalities are taking shape that benefit brands and users alike.

Generative AI has moved beyond creating single, flat images. Now, it can deliver layered compositions with editable text and elements. If you combine this with the ubiquity of agents, you basically have a co-creator sitting right there with you. Your co-creator understands color theory and composition, and can ideate with you on the design.

With agents now acting as co-creators, people are empowered to design whatever they dream about, in very professional ways. Ultimately, it closes the gap between an idea that you want to explore and what you end up producing—giving you more time to spend on creative judgment.

A year or so ago, having an AI agent as a co-creator felt far-fetched. Now, even more ambitious possibilities are within reach. Here, I share my thoughts about what's on the horizon at the nexus of AI and creativity.



Sami Ede

Co-Founder and Senior ML
Researcher, Leonardo.AI

Leonardo.AI is a powerful AI creative suite for generating production-ready images, videos, and print designs for marketing campaigns, websites, product visuals, storyboards, and more.

Sami is the Co-Founder of Leonardo.AI. He has also done AI research at the Fraunhofer HHI in Berlin, and before that worked as a full-stack developer at a German startup. He has a bachelors degree in computer science with a background in hardware and software.



True multimodality

Multimodality is extremely interesting to think about because it broadens the model's view on the world. The way I see it, this modern frontier has two ways in.

The first is to look at existing world models and their 3D environments. You can add more modalities to these interactive experiences, like a soundscape or other effects. We're starting to see some impressive results from startups in this space. But in the end, for now, these world models only learn from video images and computer game data. They can't capture the nuances of human interaction.

So the second angle, which may not be solvable in the short term, is more about asking: How do you capture the uncapturable? How do you feed lived experience into an AI model for more holistic multimodality? I look forward to seeing what researchers come up with in this space.

End-to-end agent workflows

Most of us use agents to transcribe meetings, send emails, and so on. But what if the agent that transcribes your creative brainstorming meeting then goes off and executes some of the ideas coming out of that meeting—or even orchestrates the whole campaign?

For example, say you discuss an idea for a new advertising campaign. Straight after the meeting, the agent comes up with a central character and storyline, generates consistently themed videos, stitches together clever transitions, and adds some sound. The human steps in for curation, iteration, and sign off—but the agents handle the bulk of the work.

“

You're not building tools anymore. Rather, you need to build entire workflows that capture the user end-to-end.”

A/B testing with digital twins

AI agents create an opportunity to build rich profiles. It's like digital twinning at scale. For example, say you have an agent that serves 200 customers. It can build user profiles for every one of these customers. The company can then use these profiles for A/B testing before they spend a single dollar or start building the actual product. It means you don't have to create personas anymore. Instead, you can test ideas at scale without having to talk to people.

It also leads to powerful personalization and the curation of content for individual customers. Interestingly, this might see the current flood of content abate, replaced with scarcity. Instead of throwing content out there and seeing if it sticks, companies can offer people perfectly curated experiences—something thoughtful that fits their environment and the mood they're in.

Generating spatial acoustics and haptics

AI is extremely good at learning, and it won't be long before it's creating spatial experiences. The means of distribution—like surround sound speakers and headphones—have been around for decades. If someone puts in the time, budget, and training data, I am absolutely certain that AI will be able to learn how to do this and commoditize AI-generated spatial acoustics.

Haptics is way more interesting. We already have the interface—think of a certain laptop touchpad. It may be possible to simulate textures by dragging your finger around that touchpad. Once we get more tailored actuators in this area I think haptics can easily become part of a brand's toolkit, learnable and generated by AI.



Generating relatable 3D models

The next step will be interactive worlds where users can tailor the experience in real time using text or voice input. For brands, this could replace linear go-to-market campaigns like videos or landing pages; instead offering a tailored experience within the boundaries of brand guidelines that is interactive and tailored to the user.

These 3D experiences would feel more familiar to what the user experiences day-to-day, creating more affinity and offering so much potential for brands.



What if the agent that transcribes your creative brainstorming meeting then goes off and executes some of the ideas coming out of that meeting—or even orchestrates the whole campaign?”

Advice for founders

From building tools to entire workflows

For founders wanting to pursue one of the bold new directions that generative AI is taking us, remember this: you're not building tools anymore. Rather, you need to build entire workflows that capture the user end-to-end. This way, the user never leaves your experience.

In this competitive market, you also need to consider distribution costs and compute availability. GPUs are scarce right now, and you need to factor this scarcity in. Consider whether you can guarantee enough compute for your customers as you grow. If you're offering a great product, but it's not available because the APIs are overwhelmed or you can't serve it yourself or latency is too high, then people will look to other workflows.

But when you get it right, then exciting new opportunities and pathways open up. So it's definitely a space worth getting involved in.

Build with Google's generative media models





Google develops generative media models across image, video, audio, and interactive environments, enabling entirely new ways for people to build, create, and interact with the digital world.

Image

Gemini Image

Gemini Image (Nano Banana) is Google's state-of-the-art image generation and editing model, built on Gemini. It uses advanced reasoning to deliver exceptional prompt adherence, allowing you to create and refine high-fidelity visuals conversationally.

Its most impressive capabilities include rendering accurate text directly into images, maintaining strict character and object consistency across different scenes, and grounding visuals with real-time information from Google Search.



Prompt

Editorial style photo, female model is wearing jeans, yellow top with polka dots, headband, red heels, black bag on her arm. She is holding an iced matcha latte in one hand and in the other hand she is holding a leash on a chow chow dog. She is standing in front of the house in Beverly Hills, looking into the camera. Respect the overall aesthetic and color palette of the photo with the house. There is a white logo "Love Letters" with 10% opacity shadow in the lower left corner.

[Try Gemini Image](#)[View prompt guide](#)

Video

Veo

Veo is Google's state-of-the-art video generation model, capable of producing high-fidelity 4, 6, or 8-second clips in up to 4K resolution with stunning cinematic realism.

Its most groundbreaking feature is the ability to natively generate synchronized audio—including spoken dialogue, sound effects, and ambient noise—directly alongside the video. It also offers unprecedented creative control, allowing you to seamlessly extend existing clips, use reference images to maintain strict character consistency, and precisely guide transitions by specifying the exact first and last frames of a scene.



Prompt

Camera dramatically dollies around the subject in this striking cinematic scene. It captures a high-tension moment within a long, sterile, monochromatic green corridor. A lone woman, dressed in a dark, flowing trench coat and trousers that billow dramatically, is suspended mid-air in a powerful, graceful leap. Her arms are outstretched as if bracing for impact or propelling herself forward. Her sharp profile reveals an intense, focused expression, suggesting profound determination.

[Try Veo](#)[View prompt guide](#)



Music

Lyria

Lyria is a family of music generation models designed to produce high-fidelity, professional-grade stereo audio. These models excel at structural control, allowing you to seamlessly dictate musical architecture like intros, verses, choruses, and bridges using natural language timestamps.

Lyria can generate vocal performances in eight different languages, giving you the flexibility to guide the overall vocal style and lyrics to perfectly match your composition.

Prompt

[00:00] Begin immediately with a massive gospel choir singing a powerful, uplifting harmony about being kind to yourself.

[00:15] A heavy, modern hip-hop drum beat and a deep 808 bassline drop in, matching the energy of the choir.

[00:30] A male lead vocalist begins rapping a confident verse about overcoming life's challenges, while the large choir punctuates his lines in the background.

[01:10] Transition into a huge, triumphant chorus celebrating victory and winning. The gospel choir sings at full volume, layering rich, soulful harmonies over the driving hip-hop beat and triumphant brass horns.

[01:50] The beat strips back to just a gentle Hammond B3 organ. The rapper delivers a quiet, emotional bridge about giving yourself grace, supported by soft, warm hums from the massive choir.

[02:10] The full hip-hop beat and the giant choir return at maximum energy for an uplifting final chorus, before ending on a resonant, sustained choir chord at [03:00].

[Listen here](#)[Try Lyria](#)[View prompt guide](#)

Speech

Gemini Audio

Gemini Audio is a suite of models designed to process and generate expressive sound. For audio understanding, it transforms lengthy, unstructured recordings into clean data, complete with accurate transcriptions, speaker separation, and emotion detection.

Its text-to-speech capabilities allow you to produce dynamic single or multi-speaker performances by prompting for specific styles, accents, and pacing. Additionally, it translates speech across over 70 languages while using style transfer to preserve the original speaker's unique intonation.

Prompt

[excited] Welcome to the next generation of Gemini Text-to-Speech! [medium pause] [positive] Let's skip the boring instructions and just show you what it can do. [short pause] By simply typing in brackets, you can completely hijack the emotional tone. Watch this. [short pause] We can drop down into a [slow] [whispering] dark, suspenseful whisper where every single syllable matters... [short pause] [extremely fast] [panicked] Or we can speed things up to absolute panic mode because we are running out of time! [long pause] [heavy sigh] [neutral] The model breathes. It feels. [laughs harder] It even laughs! [short pause] [positive] But honestly? My favorite feature is the pacing. [medium pause] Because sometimes, the most powerful thing you can say is... [long pause] [whispering] ...nothing at all.

[Try Gemini Audio](#)



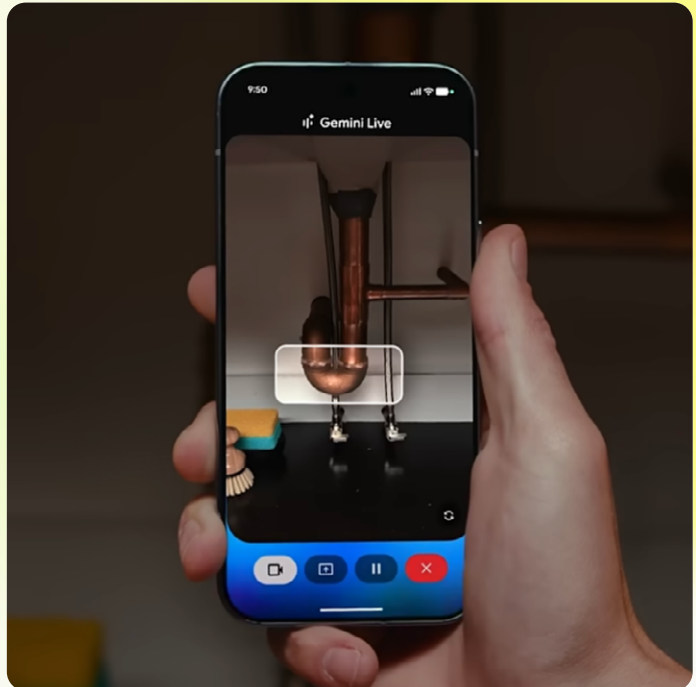
🎯 Real-time

Gemini Live

Gemini Live enables you to build low-latency voice and vision agents that interact at the natural speed of conversation. It processes continuous streams of audio, video, and text to deliver fluid spoken dialogue while effectively filtering out background noise.

The model provides robust multilingual support, seamlessly triggers external tools, and strictly adheres to complex instructions to keep your conversational agents securely on track.

[Try Gemini Live](#)



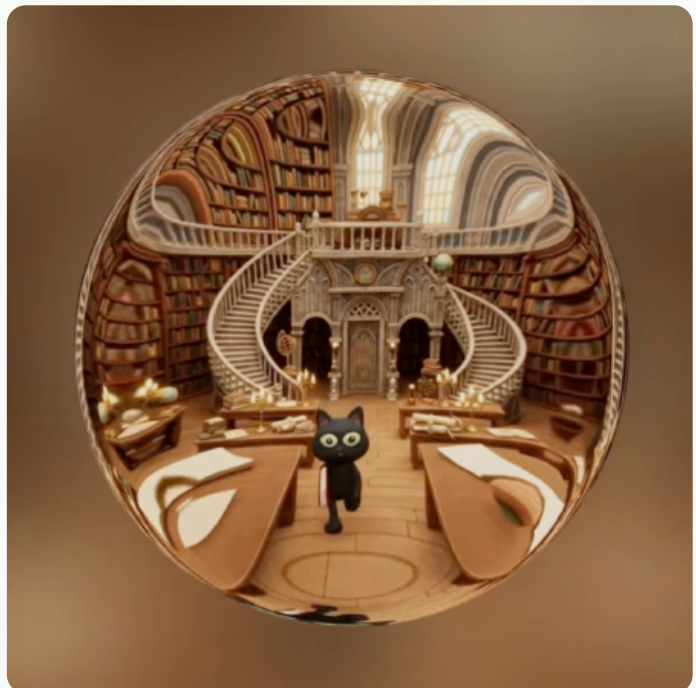
Genie

Genie is a general-purpose world model that simulates the physics and dynamics of physical environments. Rather than generating static 3D snapshots, it creates the path ahead frame-by-frame based on your text descriptions, image prompts, and in-world actions.

You can use it to sketch entirely new living ecosystems, remix existing environments, or trigger promptable events like changing weather conditions on the fly. This enables highly consistent world simulations that can be used for everything from immersive fiction to safely training autonomous vehicles.

[Try Genie](#)

[View prompt guide](#)



Google for Startups

Turn your AI vision into reality



Sign up for monthly sessions
with Google Cloud's AI experts

Sign up



Get the latest updates
with our startup newsletter

Subscribe



Contact our
startup sales team

Contact us