

What will the future of AI filmmaking be like?

The realistic, high-definition output from Sora's products is so impressive that some are even predicting that Hollywood will soon disappear. Runway's newest models can create short films that rival those made by blockbuster animation studios.

Midjourney and Stability AI are two of the most popular text-to-image models, which are also currently working on video.

Several companies are racing to gain business from these breakthroughs. "I'm constantly screaming, 'Oh my god, this is awesome' when experimenting with these tools," says Gary Lipkowitz, CEO of Vyond, a company that provides a platform for aggregating short animated videos. . 'But how can you use this at work?'

Whatever the answer, it will likely affect many businesses and change the roles of many professionals, from animators to advertisers. Concerns about misuse are also growing. The ability to create fake videos will flood the Internet with harmful content.

Here are four things to ponder about the direction of AI filmmaking.

1. Sora is just the beginning

OpenAI's Sora is currently leading the way in video creation. But other companies are working to catch up. The market will become extremely crowded over the next few months as more and more companies improve their technology and start launching products that compete with Sora.

UK-based startup Haiper revealed itself this month. It was founded in 2021 by former Google DeepMind and TikTok researchers who wanted to research a technology called neural radiation fields, or NeRF, that can convert 2D images into 3D virtual environments. They thought a tool that turned snapshots into scenes that users could walk into would be useful for creating video games.

But six months ago, Haiper pivoted from virtual environments to video clips, adapting its technology to what CEO Yishu Miao believes will be an even bigger market than gaming. Miao said: 'We realized that creating videos is a fun thing. The demand for it will be very high.'

'Air Head' is a short film made by Shy Kids, a pop band and filmmaking collective based in Toronto, with the help of Sora.

Like OpenAI's Sora, Haiper's generative video technology uses a diffusion model to manage images and a transformer (a component in large language models like GPT-4 that helps them predict what will happen very well). next), to control consistency between frames. 'Videos are sequences of data, and the transform engine is the best model to learn sequences of data,' said Miao.

Consistency is a major challenge for generative video and the main reason why existing tools only generate a few seconds of video at a time. Transformer for video creation can enhance the quality and length of clips. The downside is that the transformer creates something or causes hallucinations. In the text, this is not always clear. In the video, it can create a person with multiple heads. Maintaining the operation of the transformer requires a huge store of training data.

That's why Irreverent Labs, founded by former Microsoft researchers, is taking a different approach. Like Haiper, Irreverent Labs started out creating environments for games before moving on to full video creation. But the company doesn't want to follow the crowd by copying what OpenAI and other companies are doing. 'Because then it will be a war on computers, a total war on GPUs,' said David Raskino, co-founder and CTO of Irreverent. And there's only one winner out there."

Instead of using a transformer, Irreverent's technology combines the Diffusion model with a model that predicts what's in the next frame based on common physics, such as how a ball bounces or how water splashes. splash on the floor. Raskino said this method reduces both training costs and the number of hallucinations. The model still produces problems, he said, but they are physical distortions (such as a bouncing ball not following a smooth curve). Known mathematical fixes can be applied to the video after it is created.

Which approach will last remains to be seen. Miao compared today's technology to major language models. Five years ago, OpenAI's groundbreaking first model surprised everyone because it showed what was possible. But it will take several years for this technology to change the game. Same with the video, Miao said: 'We are all at the foot of the mountain.'

2. What will people do with generative videos?

Video is the medium of the Internet. YouTube, TikTok, newsreels, advertisements, generative videos are expected to appear everywhere there is video.

The marketing industry is one of the most enthusiastic adopters of generative technology. According to a recent survey Adobe conducted in the US, two-thirds of marketing professionals have tested Generative AI at work, with more than half saying they have used the technology to create images.

Next is generative video. Some marketing companies have made short films to demonstrate the technology's potential. The latest example is the 2.5 minute 'Somme Requiem' by Myles. You can watch the footage below in an exclusive reveal from MIT Technology Review.

'Somme Requiem' is a short film made by Los Angeles-based production company Myles. Every shot was created using Runway's Gen 2 model. The clips are then edited by a team of video editors at Myles.

'Somme Requiem' depicts snow-covered soldiers during the Christmas truce during World War I in 1914. The film is made up of dozens of different shots produced using generative video modeling from Runway, then stitched together, color graded and music added by a video editor at Myles. 'The future of storytelling will be a hybrid workflow,' said founder and CEO Josh Kahn.

Kahn chose a wartime setting to make his point. He noted that the Apple TV+ series Masters of the Air, about a group of World War II pilots, cost \$250 million. The team behind Peter Jackson's World War I documentary They Shall Not Grow Old spent four years curating and restoring more than 100 hours of archival film. 'Most filmmakers can only dream of having the opportunity to tell a story in this genre,' says Kahn.

'Independent filmmaking is almost on the verge of death,' he added. 'I think this is going to create an incredible resurgence.'

'The horror genre is where people try new things, try new things until they fail,' Raskino said. 'I think we're going to see a blockbuster horror movie created by 4 people in a basement somewhere using AI.'

So will generative video take down Hollywood? Currently, not yet. The scenes in 'Somme Requiem' - empty forests, deserted military camps - look great. But the people in there still have deformed fingers and distorted faces, characteristic of AI products. Video is at its best in wide shots or lingering close-ups, which create an eerie atmosphere but little action. If 'Somme Requiem' had gone on any longer it would have become dull.

But the background shots that appear in feature films are mostly just a few seconds long but can take hours to shoot. Raskino suggests that generative video models could soon be used to create those interlaced footage at low cost. This can also be done quickly in later stages of production without the need for reshoots.

Michal Pechoucek, CTO at Gen Digital, the cybersecurity giant behind a range of antivirus brands including Norton and Avast, agrees. 'I think this is where technology is going,' he said. 'We will see many different models, each with special training in a certain area of ??film production. These will only be tools used by talented video production teams'.

A big problem with generative video is the lack of user control over the output. Generating still images may be corrupted; Producing a few seconds of video is even more dangerous.

Miao said: 'For now, it's still very exciting, you get great moments. But making the video exactly what you want is a very difficult technical problem. Somehow, we're finding a way to create long, consistent videos from just a single prompt.'

That's why Vyond's Lipkowitz thinks the technology isn't ready for most enterprise customers yet. These users, he said, want more control over the look of their videos than what current tools give them.

Thousands of companies around the world, including approximately 65% ??of the Fortune 500, use Vyond's platform to create animated videos for internal communications, training, marketing, and more. Vyond is based on a series of generative models, including text-to-image and text-to-speech, but provides a simple drag-and-drop interface that allows users to piece videos together manually, segment by segment. one, instead create a full video clip with one click.

Running a creative model is like rolling the dice, says Lipkowitz. 'This is unlikely for most video production teams, especially in the corporate sector where everything has to be pixel-perfect and on-brand,' he says. 'Videos can turn out really bad - like characters with too many fingers or company logos in the wrong color - unfortunately, that's how the AI ??gene works'.

The solution is more data, more training, repetition. 'I wish there could be some algorithm for any problem,' Miao said. 'But no, it's all about learning more'.

3. Misinformation is not new, but deepfake will make the situation worse

Online misinformation has been undermining our trust in the media, in institutions, and in each other for years.

'We are replacing trust with distrust, confusion, fear and hate,' Pechoucek said. A society without a foundation of truth will become degenerate'.

Pechoucek is particularly worried about the malicious use of deepfakes in elections. For example, in last year's election in Slovakia, attackers shared a fake video showing the leading candidate discussing a plan to manipulate voters. The video is low quality and easily detected as a deepfake video. But Pechoucek believes that will be enough to overturn the results in favor of the other candidate.

'Adventurous Puppies' is a short clip made by OpenAI using Sora.

John Wissinger, head of strategy and innovation at Blackbird AI, a company that tracks and manages the spread of misinformation online, believes that fake videos are most convincing when they combine real and fake footage. Take two videos of President Joe Biden walking across the stage. In one place he stumbled, in another he did not. Who's to say which is real?

Says Wissinger: 'Let's say an event actually happened, but the way it was presented to me was subtly different. That might affect my emotional response to it.' As Pechoucek noted, a fake video doesn't even have to be great to make an impact. A fake video with bad intentions that fits existing biases will do more damage than an inappropriately polished product, Wissinger said.

That's why Blackbird focuses on who is sharing what with whom. In a sense, whether something is right or wrong is less important than where it comes from and how it is spread, Wissinger said. His company has been tracking low-tech misinformation, such as social media posts that show real images out of context. Generative technologies make things worse, he said, adding that people presenting in misleading ways, whether intentionally or not, is not new.

Sharing and promoting false information on social networks will make things confusing. Just know that there is a lot of fake media out there that will sow seeds of doubt in bad faith discourse. "You can see that soon we won't be able to distinguish between what's AI-generated and what's real," Wissinger said.

4. We are facing a new online reality

Fake videos will soon appear everywhere, from misinformation campaigns, to advertising spots, to Hollywood blockbusters. So what can we do to find out what is real and what is just imagination? There are many solutions but none are truly radical.

The tech industry is solving this problem. Most generative tools try to enforce certain terms of use, such as preventing people from creating videos of public figures. However, there are ways to bypass these filters, and open source versions of the tool may come with easier policies.

Companies are also developing standards for watermarking AI-generated media and tools to detect it. But not all tools add watermarks, and watermarks can be removed from the video's metadata. No reliable detection tools exist. Even when such tools work, they become part of a cat-and-mouse game of trying to keep up with advances in the models for which they are designed.

'Spaghetti Eat Will Smith' is a short film made by OpenAI using Sora.

Online platforms like X and Facebook often get poor reviews when it comes to censorship. And we shouldn't expect these platforms to improve once the problem becomes more difficult. Miao used to work at TikTok with the task of building a censorship tool to detect uploaded videos that violate TikTok's terms of use. Even Miao is wary of what's about to happen: 'There's real danger there. Don't believe everything you see on your laptop.'

Blackbird has developed a tool called Compass, which allows you to check the authenticity of articles and social media posts. Paste a link into the tool and a large language model will generate a demo drawn from trusted online sources (these are always open to review, Wissinger says) that provide some language. scene for the linked document. The results are very similar to the community notes that are sometimes attached to controversial posts on sites like X, Facebook, and Instagram.

While many people link to a fact-checking site, many others may not know such tools exist or trust them. Misinformation also tends to spread further than any subsequent correction.

Pechoucek said tech companies need to expand their software to enable more competition on safety and trust issues. That would also allow cybersecurity companies to develop third-party software to monitor this technology. That's what happened 30 years ago when Windows had problems with malware, he said: 'Microsoft allowed antivirus companies to step in to help protect Windows. As a result, the online world becomes a safer place.'

But Pechoucek is not so optimistic. "Technology developers need to build their tools with safety as the primary goal," he said. But more people think about how to make this technology more powerful than they worry about how to make it more secure.'

Video created by OpenAI using Sora.

There's a common fatalistic refrain in the tech industry: Change is coming, deal with it. "I don't think tech companies can shoulder all the responsibility," Raskino said. After all, the best defense against any technology is good education for everyone. There are no shortcuts.'

Miao agrees: 'It is inevitable that we will widely adopt generative technology. But it is also the responsibility of the entire society. We need to educate people'.

He added: 'Technology will move forward and we need to prepare for this change.' 'We need to remind our parents and friends that what they see on the screen may not be real. This is especially true for the older generations.' 'Our parents need to be aware of this danger. I think everyone should work together.'

We need to work together quickly. When Sora launched a month ago, the tech world was stunned by the rapid growth of video. But the vast majority of people don't know this type of technology even exists, Wissinger said: 'They certainly don't understand the trend we're pursuing. I think it will take the world by storm.'

You finished reading the article "**What will the future of AI filmmaking be like?**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.