

All CCS and Cross-sectoral Visual Arts Media & Press Audiovisual - Film, TV & Radio Video Games & Multi-media Tools & Resources Spotlight International News Featured Stories

OpenAI Unveils Revolutionary Text-to-Video Model 'Sora'

OpenAI's new 'Sora' model can bring science fiction movie scenes to life without a Hollywood budget. And that means that disruptive technology is now an issue the Creative Industries will need to address faster than expected.

Very soon, a scene where an astronaut will be exploring a deserted lunar landscape, will not require expensive studio production. It will cost almost nothing more than can be done in a home studio. Artificial intelligence research firm OpenAI today revealed details of a new text-to-video model called "Sora" that can generate photorealistic video scenes from simple text prompts. The results are impressive.

Using a transformer architecture, Sora can gradually remove noise and create videos of up to 60 seconds directly from text instructions, doing away with the frame-by-frame approach of past systems. This allows it to maintain visual quality and adhere closely to user prompts when animating multiple characters with complex motions and settings.

Sora can generate complex scenes with multiple characters, specific types of motion, and accurate details of the subject and background. The model understands not only what the user has asked for in the prompt, but also how those things exist in the physical world. Videos output by Sora include intricate camera movements like pans and zooms

Our introductory example would be produced along these lines: Prompt: "Show me a solitary astronaut walking across the dusty surface of the moon, leaving footprints in the untouched soil. Capture the vastness and silence of space, with only the Earth hanging like a blue marble in the distance. Hints of alien structures hidden in the shadows add a touch of mystery."

The model builds on OpenAI's prior work with DALL-E and GPT by employing a "recaptioning" technique to generate highly descriptive text for visual training data. Sora operates using a transformer architecture, which allows it to gradually remove noise and create videos in a single pass, rather than frame by frame. Sora learns 3D geometry, and consistency and even tells stories through camera angles without being explicitly programmed to do so.

The model can generate videos up to 60 seconds long, featuring complex camera motion and multiple characters with vibrant emotions. However, Sora still struggles with accurately simulating physics for complex scenes and specific cause-and-effect relationships.

Currently in "red-teaming" testing with outside experts, Sora breaks down videos into small "patches" for training rather than using searchable internet data. OpenAI Chief Scientist Dario Amodei said the next step is developing tools to detect misleading content before a future public release. Open AI is developing image classifiers to review the frames of every video generated.

Disruption and limitations

While impressive, Sora has limitations around spatial relationships, events over time and confusing visual details. For example, Sora may not fully understand cause-and-effect relationships in certain scenarios, leading to inaccurate representations. Also, it may have difficulty maintaining spatial consistency and accurately representing left-right relationships in scenes.

Sora can generate high-quality video content much faster than traditional methods, making it easier for professionals to create and distribute content. Sora's ability to create cinematic videos with an emergent understanding of cinematic grammar can enhance storytelling skills. Major commercial applications will include using Sora to create video advertisements, explainer videos for products/services, video tutorials, celebrity deepfakes for entertainment, and custom videos for social media engagement.

While impressive, Sora is still in the "red-teaming" testing phase and is not publicly available yet. It's being tested to prevent

harmful outputs. A select group of visual artists, designers and filmmakers are participating in the testing process to provide feedback. The goal is to ensure Sora enhances creative work rather than replacing professionals. Its full impact on industries won't be clear until it's released publicly and used by businesses. Existing demos show Sora's potential applications, but no timeline is given for public release.

The ability to create highly realistic videos raises questions about authenticity, media literacy and the need for regulation. Sora is expected to have a significant impact on video production by streamlining the workflow for professionals in the creative industries, resulting in the creation of video content at a much faster pace than traditional methods. It is expected to affect the Advertising industry, the Gaming industry, the music video production industry and even film production.

Major applications will be casual video creation for social media and speeding up game/film pre-production workflows. Sora dramatically lowers barriers to video making but cannot replace all aspects of professional filmmaking.

Commercial applications could include the following:

Video advertising: Using Sora to automatically generate customized and personalized video ads.

Explainer videos: Creating explanatory videos about products, services, processes, etc. for marketing and educational purposes.
Video tutorials: Generating instructional tutorial videos across different topics easily.
Entertainment: Potential uses in movie/TV special effects, virtual avatars, and celebrity deepfakes.
Social media: Empower individual content creators to generate high-quality video clips for platforms like TikTok, Instagram, YouTube, etc.
Video games: Speeding up development cycles by utilizing Sora to generate cinematics, cutscenes, and visual assets automatically.
Film/TV production: Initial uses in concept/storyboarding before full production, with potential overtime for automated low-budget content.
Education: Generating personalized video lessons, simulations and virtual demonstrations at scale for distance learning.

Read further
OpenAI's Sora Turns AI Prompts Into Photorealistic Videos

[What is Sora and how does it work? A guide to OpenAI's latest text-to-video AI tool]