

## Meet DALL-E, the A.I. That Draws Anything at Your Command

New technology that blends language and images could serve graphic artists — and speed disinformation campaigns.



By Cade Metz

April 6, 2022

SAN FRANCISCO — At OpenAI, one of the world’s most ambitious artificial intelligence labs, researchers are building technology that lets you create digital images simply by describing what you want to see.

They call it DALL-E in a nod to both “WALL-E,” the 2008 animated movie about an autonomous robot, and Salvador Dalí, the surrealist painter.

OpenAI, backed by a billion dollars in funding from Microsoft, is not yet sharing the technology with the general public. But on a recent afternoon, Alex Nichol, one of the researchers behind the system, demonstrated how it works.

When he asked for “a teapot in the shape of an avocado,” typing those words into a largely empty computer screen, the system created 10 distinct images of a dark green avocado teapot, some with pits and some without. “DALL-E is good at avocados,” Mr. Nichol said.



Image Generated by AI

DALL-E generated this image from a command for “cats playing chess.” OpenAI

When he typed “cats playing chess,” it put two fluffy kittens on either side of a checkered game board, 32 chess pieces lined up between them. When he summoned “a teddy bear playing a trumpet underwater,” one image showed tiny air bubbles rising from the end of the bear’s trumpet toward the surface of the water.

DALL-E can also edit photos. When Mr. Nichol erased the teddy bear’s trumpet and asked for a guitar instead, a guitar appeared between the furry arms.

A team of seven researchers spent two years developing the technology, which OpenAI plans to eventually offer as a tool for people like graphic artists, providing new shortcuts and new ideas as they create and edit digital images. Computer programmers already use Copilot, a tool based on similar technology from OpenAI, to generate snippets of software code.

But for many experts, DALL-E is worrisome. As this kind of technology continues to improve, they say, it could help spread disinformation across the internet, feeding the kind of online campaigns that may have helped sway the 2016 presidential election.

“You could use it for good things, but certainly you could use it for all sorts of other crazy, worrying applications, and that includes deep fakes,” like misleading photos and videos, said Subbarao Kambhampati, a professor of computer science at Arizona State University.

A half decade ago, the world’s leading A.I. labs built systems that could identify objects in digital images and even generate images on their own, including flowers, dogs, cars and faces. A few years later, they built systems that could do much the same with written language, summarizing articles, answering questions, generating tweets and even writing blog posts.

Now, researchers are combining those technologies to create new forms of A.I. DALL-E is a notable step forward because it juggles both language and images and, in some cases, grasps the relationship between the two.

“We can now use multiple, intersecting streams of information to create better and better technology,” said Oren Etzioni, chief executive of the Allen Institute for Artificial Intelligence, an artificial intelligence lab in Seattle.

The technology is not perfect. When Mr. Nichol asked DALL-E to “put the Eiffel Tower on the moon,” it did not quite grasp the idea. It put the moon in the sky above the tower. When he asked for “a living room filled with sand,” it produced a scene that looked more like a construction site than a living room.

But when Mr. Nichol tweaked his requests a little, adding or subtracting a few words here or there, it provided what he wanted. When he asked for “a piano in a living room filled with sand,” the image looked more like a beach in a living room.



DALL-E generated this image when asked for “a living room filled with sand, sand on the floor, piano in the room.” OpenAI

DALL-E is what artificial intelligence researchers call a neural network, which is a mathematical system loosely modeled on the network of neurons in the brain. That is the same technology that recognizes the commands spoken into smartphones and identifies the presence of pedestrians as self-driving cars navigate city streets.

A neural network learns skills by analyzing large amounts of data. By pinpointing patterns in thousands of avocado photos, for example, it can learn to recognize an avocado. DALL-E looks for patterns as it analyzes millions of digital images as well as text captions that describe what each image depicts. In this way, it learns to recognize the links between the images and the words.

When someone describes an image for DALL-E, it generates a set of key features that this image might include. One feature might be the line at the edge of a trumpet. Another might be the curve at the top of a teddy bear’s ear.

Then, a second neural network, called a diffusion model, creates the image and generates the pixels needed to realize these features. The latest version of DALL-E, unveiled on Wednesday with a new research paper describing the system, generates high-resolution images that in many cases look like photos.

Though DALL-E often fails to understand what someone has described and sometimes mangles the image it produces, OpenAI continues to improve the technology. Researchers can often refine the skills of a neural network by feeding it even larger amounts of data.

They can also build more powerful systems by applying the same concepts to new types of data. The Allen Institute recently created a system that can analyze audio as well as imagery and text. After analyzing millions of YouTube videos, including audio tracks and captions, it learned to identify particular moments in TV shows or movies, like a barking dog or a shutting door.

Experts believe researchers will continue to hone such systems. Ultimately, those systems could help companies improve search engines, digital assistants and other common technologies as well as automate new tasks for graphic artists, programmers and other professionals.

But there are caveats to that potential. The A.I. systems can show bias against women and people of color, in part because they learn their skills from enormous pools of online text, images and other data that show bias. They could be used to generate pornography, hate speech and other offensive material. And many experts believe the technology will eventually make it so easy to create disinformation, people will have to be skeptical of nearly everything they see online.

“We can forge text. We can put text into someone’s voice. And we can forge images and videos,” Dr. Etzioni said. “There is already disinformation online, but the worry is that this scale disinformation to new levels.”

OpenAI is keeping a tight leash on DALL-E. It would not let outsiders use the system on their own. It puts a watermark in the corner of each image it generates. And though the lab plans on opening the system to testers this week, the group will be small.

The system also includes filters that prevent users from generating what it deems inappropriate images. When asked for “a pig with the head of a sheep,” it declined to produce an image. The combination of the words “pig” and “head” most likely tripped OpenAI’s anti-bullying filters, according to the lab.

“This is not a product,” said Mira Murati, OpenAI’s head of research. “The idea is understand capabilities and limitations and give us the opportunity to build in mitigation.”

OpenAI can control the system’s behavior in some ways. But others across the globe may soon create similar technology that puts the same powers in the hands of almost anyone. Working from a research paper describing an early version of DALL-E, Boris Dayma, an independent researcher in Houston, has already built and released a simpler version of the technology.

“People need to know that the images they see may not be real,” he said.

Cade Metz is a technology correspondent, covering artificial intelligence, driverless cars, robotics, virtual reality, and other emerging areas. He previously wrote for Wired magazine. @cademetz

A version of this article appears in print on , Section B, Page 1 of the New York edition with the headline: An A.I. Artist Draws at Your Command