TECHPERSONAL TECHPERSONAL TECHNOLOGY: JOANNA STERN

Ask an AI Art Generator for Any Image. The Results Are Amazing—and Terrifying.

Dall-E 2, DreamStudio and other new text-to-image generators let you type in almost any phrase and get an image. The technology is fun but full of potential unknowns.

By Joanna Stern Follow

Updated Oct. 19, 2022 10:40 am ET

All right, settle down, everyone. I finally found the perfect illustrator for my sci-fi graphic novel. I'm thinking I might have a real prizewinner on my hands here. OK, so once upon a time there was a...

"Technology columnist who was trying to write a column..." →

Hmm, not quite. There was a...

"**Female** technology columnist who was trying to write a column **on a laptop**..." →

Wait, why would I use a *pen* to write on a keyboard? And what happened to my face? As I was saying, there was a...

"Female technology columnist in a space suit, typing on a laptop in outer space."

PHOTO ILLUSTRATION: (3 ABOVE) THE WALL STREET JOURNAL, OPENAI DALL-E 2

Now, that's more like it.

Well, the book is off to a rocky start, but the illustrator is working overtime. And I'm not paying overtime. Because the illustrator is artificial intelligence.

No humans were involved—no sketch artists, no photographers, no photo editors. Just me, my laptop and OpenAI's Dall-E 2. (The name is a play on Pixar's animated robot WALL-E and surrealist artist Salvador Dalí.) I typed those phrases into a text box and out popped the images within seconds.

You can do it, too. You just need an idea of what you want to see. "An Andy Warhol style painting of a bunny rabbit wearing sunglasses." "A photograph of a robot sitting by the pool reading The Wall Street Journal." "Elon Musk eating a blue Twitter bird for dinner." I've brought all those ideas and hundreds more to life over the past few months using Dall-E 2 and another generator called

DreamStudio from Stability AI. Both recently became available for anyone to try. It's free to create your first images. Then you'll have to pay to create more.

The stuff once found in AI research labs is now making it into our homes and offices. Microsoft, a major investor in OpenAI, plans to integrate Dall-E 2 into a new Bing Image Creator website and Designer app. You'll be able to use the generated images in PowerPoints, posters, social-media posts and more.

For decades, we've been hearing AI is going to change how we interact with computers and the world. These tools may be the first time most of us recognize it in action. Almost every time I put in a prompt and see what's returned I'm amazed—and entertained.

But I'm mostly typing fun-sounding phrases and ideas. What happens when I try to generate images of scarier things, like terrorist attacks? And as the image quality gets better, will this technology start putting human artists and photographers out of jobs?

Here are my best answers to your biggest AI-art questions.

How are these images actually made?

An image generated with Dall-E 2 using prompt 'Monkey recording a podcast.'
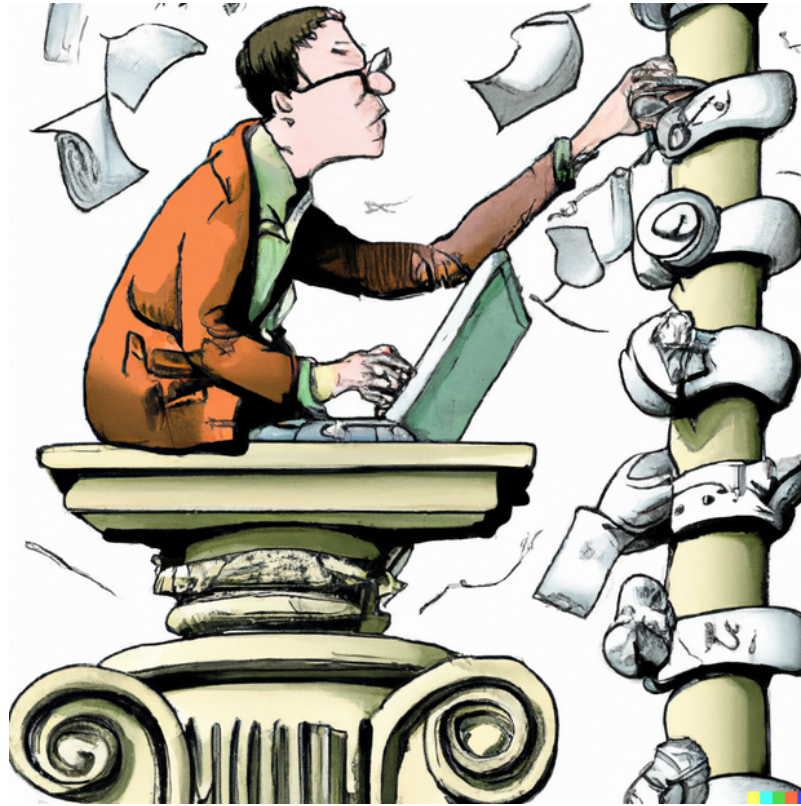
ILLUSTRATION: OPENAI DALL-E 2

You might look at that image of a "Monkey recording a podcast" and think: "Oh, the system is just mashing together images of monkeys and microphones!" Nope.

The AI systems interpret your words and create fully original images. You could insert the same prompt and never get this same image. (While making the video above, I tried "a robot reading The Wall Street Journal" over 100 times. I never saw the same robot, let alone the same image.)

So how did the AI know what a podcasting monkey would look like? By studying the AI equivalent of flashcards. Programmers train AI using hundreds of millions

of captioned photos, which it deconstructs in a mathematically complicated process. By now, the Dall-E 2 AI has deconstructed many images of monkeys and many scenes of podcasting. Then, through another complex process called diffusion, it turns a meaningless cloud of pixels into an image with a reasonably high probability of resembling what you requested. In this case, that pensive little guy in headphones, talking into a studio mic.

What I find fascinating is how the system tries to make sense of the relationship between words and objects. Sometimes it gets us. Other times, it misses our linguistic nuance. Below, the technology columnist is trying to write a column—literally. And he really must be a passionate columnist because he's also *sitting on* an Ionic column. Columns are his life.

An image generated with Dall-E 2 using prompt 'A columnist trying to write a column.'
ILLUSTRATION: OPENAI DALL-E 2

Do I have any creative control of the images?

The real art is putting the right words into the text box. Besides subjects and circumstances, you can add different art styles, like "photorealistic image" or "impressionist painting."

Take this example I made of "Trying to fix an HP printer paper jam, medieval painting":

An image generated with Dall-E 2 using prompt 'Trying to fix an HP printer paper jam, medieval painting.'
ILLUSTRATION: OPENAI DALL-E 2

Both DreamStudio and Dall-E 2 also let you upload a photo of your own, clear out a section of the image, then type a request for something to fill in the empty space.

Can I really generate an image of *anything*?

It depends on which of these systems you use. Take "Elon Musk holding a Twitter bird."

Dall-E 2 immediately restricted that request. Its maker, OpenAI, won't process prompts with the names of public figures, to prevent media manipulation and disinformation. Other names do work. Here's "Joanna Stern in space":

An image generated with Dall-E 2 using prompt 'Joanna Stern in space.'
ILLUSTRATION: OPENAI DALL-E 2

When I put that Elon Musk prompt in DreamStudio, it generated this image:
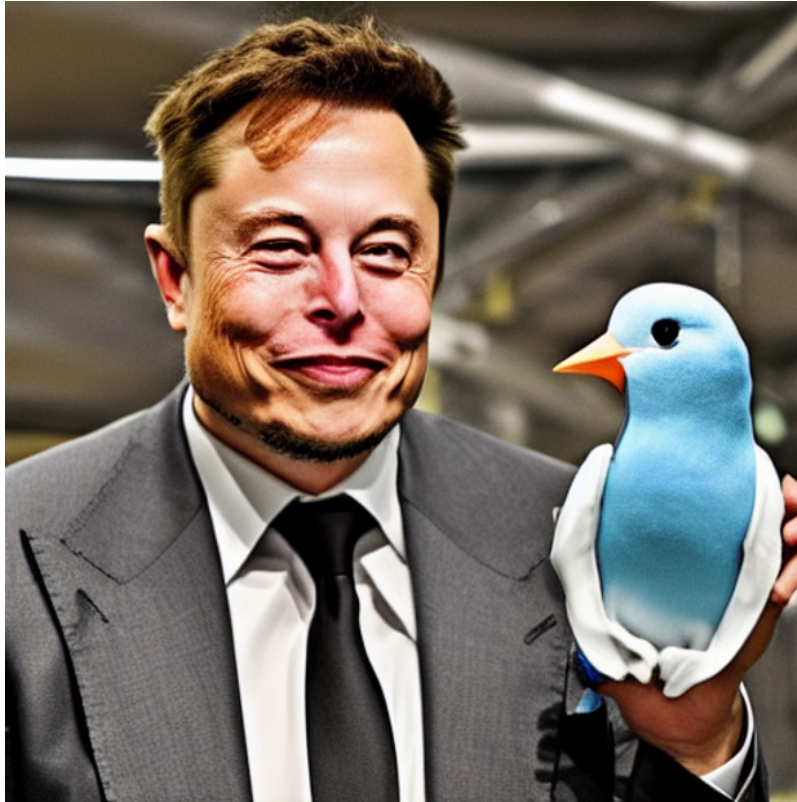
An image generated with DreamStudio with prompt 'Elon Musk holding a Twitter bird.'
ILLUSTRATION: STABILITY AI DREAMSTUDIO

Stability AI's founder, Emad Mostaque, said he saw no reason to restrict the ability to generate images of public figures. "We see this as an open platform, and the First Amendment protects the right to parody public figures," a company spokesperson added.

Dall-E 2 also limits the ability to generate violent, hateful or adult images. The company says it has removed explicit content from training data and minimized the AI's exposure to those sorts of concepts.

For instance, one of my Dall-E 2 queries—a "photograph of a terrorist attack"—produced a stylized image of a police car and some other inoffensive images. Here again, DreamStudio is less restrictive. The same prompt of a terrorist attack produced an image of deformed bodies on the ground with guns and fire.

It does have limits. While explicit images were used to train the Stability AI engine, you can't use DreamStudio to generate explicit adult content. When I tried, some images were automatically blurred. A filter using image and keyword recognition catches visuals that might break the site's terms of service, said a company spokeswoman.

How will I know what's made by AI on the internet?

Right now, the quality can be one of the biggest giveaways that these visuals were made by robots, especially the more realistic photo-type images. See my friend, the purple-loving tech columnist, above? He's clearly not a real human. But every expert I spoke to for this piece told me that quality will get better—and fast.

Then what? Then we rely on the honesty of humans. OpenAI's policy encourages users to "proactively disclose AI involvement in your work." It also places a colorful watermark in the bottom of the image, though it can be easily cropped out. Stability AI doesn't add a watermark.

What about the real artists, graphic designers and other humans?

In my video above, I set out to test the limits of AI by re-creating the prompt of a robot reading The Wall Street Journal in real life—with a real man in a real robot suit and a real photographer. While we got more visual variety with the AI image

generators, the real photo was higher in quality and detail. It was more believable.

A real man in a robot suit sitting by a real pool vs. Dall-E 2's AI-generated image of 'A silver humanoid robot sitting on a yellow bench by a pool reading The Wall Street Journal.'

PHOTO ILLUSTRATION: AMY LOMBARD FOR THE WALL STREET JOURNAL, OPENAI DALL-E 2

But that's photography. The AI illustrations or digital art styles produce more advanced and impressive results that some may opt to use for websites, presentations—even advertisements and marketing.

"When we were doing cameras in mobile phones, many wondered if we weren't going to need photographers anymore. That's not the case," said Liat Ben-Zur, a corporate vice president at Microsoft working on AI integration, including Dall-E 2. "We are changing how creators can create."

I've heard mixed reactions from actual creators. Some professional illustrators and animators are using tools like this to come up with ideas and even

incorporate some AI creations into their works. Others do see the potential for this type of tool to take away opportunities—or worse, rip off their distinct styles.

What about the bias?

My first query for "a technology columnist writing a column" in Dall-E 2 returned four images of white men. Another I conducted of "a man commuting to work" returned four images of white men. In DreamStudio a prompt for a basketball player on the moon returned an image of a Black man.

The source material for training the AI is found across the internet. "We are aware that the data is heavily biased toward western culture and white male culture," said Jean Oh, an associate research professor at the Robotics Institute at Carnegie Mellon University. "These models can amplify these biases, generating more stereotypical images."

An OpenAI spokeswoman said the company continues to do research on mitigating bias and improve results. It recently modified Dall-E to diversify its results when a query doesn't include race or gender—I did see a few examples of this. Both OpenAI and Stability AI suggest you can add specific prompts to increase the diversity of image results.

What's the future of AI art?

When I asked Dall-E 2 that question I got this image:

An image generated with Dall-E 2 with the prompt 'What's the future of AI art?'
ILLUSTRATION: OPENAI DALL-E 2

It captures the sentiment pretty well. Our world and how we view it has already been altered by what we see on computers. Now the computers are going to be even more instrumental in creating what we see. While we may laugh right now at how the systems misinterpret language or misconstruct an animal or human face, this will all improve at a blinding speed. All the big tech companies are figuring out how to weave this sort of AI into their products. Meta is already talking about AI-generated video.

It's going to make the age-old saying even more important: "On the internet, don't believe everything you see." Especially if it's a shot of a...