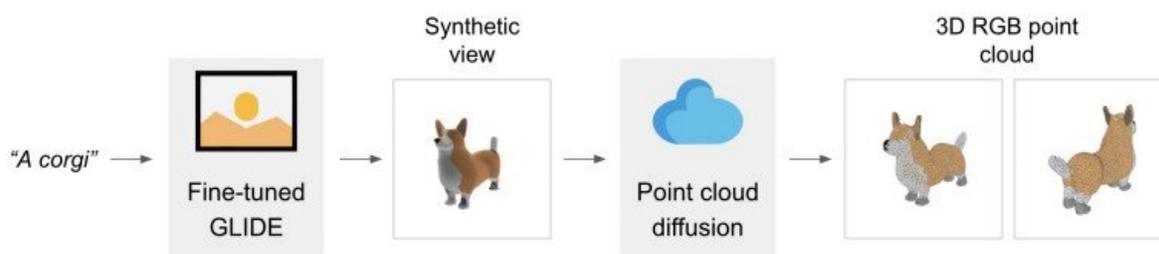


# OpenAI announces Point-E, a machine learning system that quickly creates 3D images from a text prompt

December 21 2022, by Bob Yirka

---



A high-level overview of the pipeline. First, a text prompt is fed into a GLIDE model to produce a synthetic rendered view. Next, a point cloud diffusion stack conditions on this image to produce a 3D RGB point cloud. Credit: *arXiv* (2022). DOI: 10.48550/arxiv.2212.08751

A team of researchers at San Francisco-based OpenAI, has announced the development of a machine-learning system that can create 3D images from text much more quickly than other systems. The group has published a paper describing their new system, called Point-E, on the *arXiv* preprint server.

Over the past year, several groups have announced products or systems that can generate a 3D-modeled image based on a text prompt, e.g., "a blue chair on a red floor," or "a young boy wearing a green hat and riding a purple bicycle." Such systems generally have two parts. The first

reads the text and tries to make sense of it. The second, trained on internet searches, renders the desired image.

Because of the complexity of the task, these systems can take a long time to return a model, ranging from hours to days. In this new effort, the researchers built a similar system that returns results within minutes, though they readily acknowledge that the results "fall short of the state-of-the-art in terms of sample quality."

To create images more quickly, the researchers adopted an approach somewhat different than others. Their system does not even create images in the traditional sense. Instead, it generates point clouds, which, when viewed together, resemble the desired image. The team took this approach because generating point clouds is far easier than generating actual images. To create the results, the system routes images it finds through another AI system they developed that converts what it receives to meshes, which produce the 3D point cloud model of the intended object.

The first part of the system was made using two modules—the first converts the text into an image idea and the second part finds images that are used to generate a generic image. In operation, the system runs very much the same as others of its kind—a user inputs a descriptive text prompt and the system returns an image model. They note that while the [visual quality](#) is not comparable to other systems, it might be more suitable to other applications, such as fabricating real-world objects via a 3D printer.

The researchers have made the system [open access](#)—users who wish to work with it can access the code on [GitHub](#).

**More information:** Alex Nichol et al, Point-E: A System for Generating 3D Point Clouds from Complex Prompts, *arXiv* (2022).

[DOI: 10.48550/arxiv.2212.08751](https://doi.org/10.48550/arxiv.2212.08751)

© 2022 Science X Network

Citation: OpenAI announces Point-E, a machine learning system that quickly creates 3D images from a text prompt (2022, December 21) retrieved 26 April 2024 from <https://techxplore.com/news/2022-12-openai-point-e-machine-quickly-3d.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.