

Generative AI explained

November 9 2023, by Adam Zewe



What do people mean when they say "generative AI," and why do these systems seem to be finding their way into practically every application imaginable? MIT AI experts help break down the ins and outs of this increasingly popular, and ubiquitous, technology. Credit: Jose-Luis Olivares/MIT

A quick scan of the headlines makes it seem like generative artificial intelligence is everywhere these days. In fact, some of those headlines

may actually have been written by generative AI, like OpenAI's ChatGPT, a chatbot that has demonstrated an uncanny ability to produce text that seems to have been written by a human.

But what do people really mean when they say "generative AI?"

Before the generative AI boom of the past few years, when people talked about AI, typically they were talking about machine-learning models that can learn to make a prediction based on data. For instance, such models are trained, using millions of examples, to predict whether a certain X-ray shows signs of a tumor or if a particular borrower is likely to default on a loan.

Generative AI can be thought of as a [machine-learning model](#) that is trained to create new data, rather than making a prediction about a specific dataset. A generative AI system is one that learns to generate more objects that look like the data it was trained on.

"When it comes to the actual machinery underlying generative AI and other types of AI, the distinctions can be a little bit blurry. Oftentimes, the same algorithms can be used for both," says Phillip Isola, an associate professor of electrical engineering and computer science at MIT, and a member of the Computer Science and Artificial Intelligence Laboratory (CSAIL).

And despite the hype that came with the release of ChatGPT and its counterparts, the technology itself isn't brand new. These powerful [machine-learning models](#) draw on research and computational advances that go back more than 50 years.

An increase in complexity

An early example of generative AI is a much simpler model known as a

Markov chain. The technique is named for Andrey Markov, a Russian mathematician who in 1906 introduced this statistical method to model the behavior of random processes. In machine learning, Markov models have long been used for next-word prediction tasks, like the autocomplete function in an email program.

In text prediction, a Markov model generates the next word in a sentence by looking at the previous word or a few previous words. But because these simple models can only look back that far, they aren't good at generating plausible text, says Tommi Jaakkola, the Thomas Siebel Professor of Electrical Engineering and Computer Science at MIT, who is also a member of CSAIL and the Institute for Data, Systems, and Society (IDSS).

"We were generating things way before the last decade, but the major distinction here is in terms of the complexity of objects we can generate and the scale at which we can train these models," he explains.

Just a few years ago, researchers tended to focus on finding a machine-learning algorithm that makes the best use of a specific dataset. But that focus has shifted a bit, and many researchers are now using larger datasets, perhaps with hundreds of millions or even billions of [data points](#), to train models that can achieve impressive results.

The base models underlying ChatGPT and similar systems work in much the same way as a Markov model. But one big difference is that ChatGPT is far larger and more complex, with billions of parameters. And it has been trained on an enormous amount of data—in this case, much of the publicly available text on the internet.

In this huge corpus of text, words and sentences appear in sequences with certain dependencies. This recurrence helps the model understand how to cut text into statistical chunks that have some predictability. It

learns the patterns of these blocks of text and uses this knowledge to propose what might come next.

More powerful architectures

While bigger datasets are one catalyst that led to the generative AI boom, a variety of major research advances also led to more complex deep-learning architectures.

In 2014, a machine-learning architecture known as a generative adversarial network (GAN) was proposed by researchers at the University of Montreal. GANs use two models that work in tandem: One learns to generate a target output (like an image) and the other learns to discriminate true data from the generator's output. The generator tries to fool the discriminator, and in the process learns to make more realistic outputs. The image generator StyleGAN is based on these types of models.

Diffusion models were introduced a year later by researchers at Stanford University and the University of California at Berkeley. By iteratively refining their output, these models learn to generate new data samples that resemble samples in a training dataset, and have been used to create realistic-looking images. A diffusion model is at the heart of the text-to-image generation system Stable Diffusion.

In 2017, researchers at Google introduced the transformer architecture, which has been used to develop large language models, like those that power ChatGPT. In [natural language processing](#), a transformer encodes each word in a corpus of text as a token and then generates an attention map, which captures each token's relationships with all other tokens. This attention map helps the transformer understand context when it generates new text.

These are only a few of many approaches that can be used for generative AI.

A range of applications

What all of these approaches have in common is that they convert inputs into a set of tokens, which are numerical representations of chunks of data. As long as your data can be converted into this standard, token format, then in theory, you could apply these methods to generate new data that look similar.

"Your mileage might vary, depending on how noisy your data are and how difficult the signal is to extract, but it is really getting closer to the way a general-purpose CPU can take in any kind of data and start processing it in a unified way," Isola says.

This opens up a huge array of applications for generative AI.

For instance, Isola's group is using generative AI to create synthetic image data that could be used to train another intelligent system, such as by teaching a computer vision model how to recognize objects.

Jaakkola's group is using generative AI to design novel protein structures or valid crystal structures that specify new materials. The same way a generative model learns the dependencies of language, if it's shown crystal structures instead, it can learn the relationships that make structures stable and realizable, he explains.

But while generative models can achieve incredible results, they aren't the best choice for all types of data. For tasks that involve making predictions on structured data, like the tabular data in a spreadsheet, generative AI models tend to be outperformed by traditional [machine-learning](#) methods, says Devavrat Shah, the Andrew and Erna Viterbi

Professor in Electrical Engineering and Computer Science at MIT and a member of IDSS and of the Laboratory for Information and Decision Systems.

"The highest value they have, in my mind, is to become this terrific interface to machines that are human friendly. Previously, humans had to talk to machines in the language of machines to make things happen. Now, this interface has figured out how to talk to both humans and machines," says Shah.

Raising red flags

Generative AI chatbots are now being used in call centers to field questions from human customers, but this application underscores one potential red flag of implementing these models—worker displacement.

In addition, generative AI can inherit and proliferate biases that exist in training data, or amplify hate speech and false statements. The models have the capacity to plagiarize, and can generate content that looks like it was produced by a specific human creator, raising potential copyright issues.

On the other side, Shah proposes that generative AI could empower artists, who could use generative tools to help them make creative content they might not otherwise have the means to produce.

In the future, he sees generative AI changing the economics in many disciplines.

One promising future direction Isola sees for generative AI is its use for fabrication. Instead of having a [model](#) make an image of a chair, perhaps it could generate a plan for a chair that could be produced.

He also sees future uses for generative AI systems in developing more generally intelligent AI agents.

"There are differences in how these models work and how we think the human brain works, but I think there are also similarities. We have the ability to think and dream in our heads, to come up with interesting ideas or plans, and I think generative AI is one of the tools that will empower agents to do that, as well," Isola says.

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: Generative AI explained (2023, November 9) retrieved 12 May 2024 from <https://techxplore.com/news/2023-11-generative-ai.html>

<p>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.</p>
--