

“Language Models are Few-Shot Learners”

GPT-3 is a powerful language model, the result of work by our paper’s 31 authors and many others at OpenAI and elsewhere who provided support. GPT-3 represents a significant shift from AI systems that rely on humans (via researchers) specifying training algorithms, to AI systems that are empowered to directly interact with humans using natural language.

We set out to make a language model 10 times larger than any previous model, and out of this model emerged a new capability: “few-shot” learning—which enables a wide range of tasks simply by asking the model in words, or by providing a few examples, instead of “fine-tuning” with thousands of examples. This lends itself to a range of practical applications including semantic search and conversational question-answering.

As with all increasingly powerful generative models, fairness and misuse are important concerns. We studied biases in GPT-3 around race, gender, and religion as a first step towards efforts to reduce these biases, and our follow-up work is ongoing.

GPT-3 itself has a number of limitations. It isn’t good at some tasks, it is text-only, and we are probably approaching the limits of scaling. It will need to be combined with new ideas to address those issues, as well as the critical ethical issues mentioned above.

OpenAI Media Contact:

Niko Felix
niko@theoutcastagency.com

About the OpenAI Authors:

Tom Brown is a researcher at OpenAI who works on efficient scaling of neural networks.

Benjamin Mann is a software engineer at OpenAI working on large scale data and infrastructure.

Nick Ryder is a researcher at OpenAI who focuses on systems and empirical scaling for large language models.

Melanie Subbiah worked on GPT-3 at OpenAI before beginning a PhD in NLP at Columbia University this year.

Jared Kaplan is a Professor in the Physics and Astronomy Department at Johns Hopkins University and a research consultant with OpenAI.

Prafulla Dhariwal is a research scientist at OpenAI working on large scale generative models and unsupervised learning.

Arvind Neelakantan is a researcher at OpenAI who works on deep learning, NLP, and knowledge reasoning.

Pranav Shyam is a researcher at OpenAI studying Deep and Unsupervised Learning.

Girish Sastry is a policy researcher at OpenAI who works on evaluating AI capabilities and preventing AI misuse.

Amanda Askell is a research scientist at OpenAI who currently focuses on evaluating AI systems and constructing human baselines for AI tasks.

Sandhini Agarwal is a policy researcher at OpenAI who works on fairness, bias and social safety.

Ariel Herbert-Voss is a research scientist at OpenAI and PhD student at Harvard University working on securing AI systems and threat intel to mitigate misuse.

Gretchen Krueger is the policy manager at OpenAI and works on model documentation and mechanisms to support responsible AI development.

Tom Henighan is a research engineer at OpenAI who works on generative modeling

Rewon Child researches large-scale generative modeling, and worked on the Algorithms team at OpenAI.

Aditya Ramesh is a researcher at OpenAI who works on grounding language in vision and large-scale generative modeling.

Daniel Ziegler is a research engineer at OpenAI who builds mechanisms enabling humans to align powerful AI systems.

Jeffrey Wu is a researcher at OpenAI who works on fine-tuning language models to learn from humans.

Clemens Winter is a research engineer at OpenAI who works on neural network optimizers and distributed systems.

Christopher Hesse is an engineer at OpenAI working on research tools.

Mark Chen is a research scientist at OpenAI working on unsupervised learning and generative modeling at scale.

Eric Sigler is a software engineer at OpenAI who works on supercomputing infrastructure for large scale AI models.

Mateusz Litwin is a software engineer at OpenAI who works on supercomputing infrastructure for large scale AI models.

Scott Gray is a research engineer at OpenAI who works on low level software optimization of AI compute.

Benjamin Chess is the engineering manager for Supercomputing at OpenAI

Jack Clark is the policy director of OpenAI, and co-chair of the AI Index at Stanford University

Christopher Berner is the Head of Compute at OpenAI

Sam McCandlish leads a team at OpenAI studying the scientific principles that underlie ML performance and scaling.

Alec Radford is a researcher at OpenAI focusing on generative models who developed DCGAN and GPT.

Ilya Sutskever is a co-founder and the Chief Scientist of OpenAI.

Dario Amodei is Vice President of Research at OpenAI and works on large language models and safety techniques.



Most of the OpenAI Authors