Generative AI – References

Uwe Pleban, Ph.D.

Last Update: 14. July 2023

References

→ A. Vasvani et al. (Google) – "*Attention Is All You Need*", [1706.03762] Attention Is All You Need (arxiv.org), June 2017

  We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train.

A. Radford et al. (OpenAI) – "*Improving Language Understanding by Generative Pre-Training*", https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf, 2018 (GPT paper)

  In this paper, we explore a semi-supervised approach for language understanding tasks using a combination of unsupervised pre-training and supervised fine-tuning. Our goal is to learn a universal representation that transfers with little adaptation to a wide range of tasks. We assume access to a large corpus of unlabeled text and several datasets with manually annotated training examples (target tasks). Our setup does not require these target tasks to be in the same domain as the unlabeled corpus. We employ a two-stage training procedure. First, we use a language modeling objective on the unlabeled data to learn the initial parameters of a neural network model. Subsequently, we adapt these parameters to a target task using the corresponding supervised objective.

→ OpenAI Blog Entry – "*Better Language Models and Their Implications*", https://blog.openai.com/better-language-models/, February 2019 (GPT-2 Announcement)

  We've trained a large-scale unsupervised language model which generates coherent paragraphs of text, achieves state-of-the-art performance on many language modeling benchmarks, and performs rudimentary reading comprehension, machine translation, question answering, and summarization—all without task-specific training.

→ A. Radford et al. (OpenAI) – "*Language Models are Unsupervised Multitask Learners*", https://d4mucfpksywv.cloudfront.net/better-language-models/language_models_are_unsupervised_multitask_learners.pdf, February 2019 (GPT-2 paper)

→ T. Brown et al. (OpenAI) – "*Language Models are Few-Shot Learners*", https://arxiv.org/pdf/2005.14165.pdf, July 2020 (GPT-3 paper)

  Here we show that scaling up language models greatly improves task-agnostic, few-shot performance, sometimes even reaching competitiveness with prior state-of-the-art fine-tuning approaches. Specifically, we train GPT-3, an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and test its performance in the few-shot setting. … GPT-3 achieves strong performance on many NLP datasets, including translation, question-answering, and cloze tasks, as well as several tasks that require on-the-fly reasoning or domain adaptation, such as unscrambling words, using a novel word in a sentence, or performing 3-digit arithmetic.

OpenAI Blog Entry – "*OpenAI API*", OpenAI API, June 2020

  We're releasing an API for accessing new AI models developed by OpenAI. Unlike most AI systems which are designed for one use-case, the API today provides a general-purpose "text in, text out" interface, allowing users to try it on virtually any English language task.

References

OpenAI Blog Entry – "*WebGPT: Improving the Factual Accuracy of Language Models through Web Browsing*", https://openai.com/blog/webgpt/, December 2021

> We've fine-tuned GPT-3 to more accurately answer open-ended questions using a text-based web browser. Our prototype copies how humans research answers to questions online—it submits search queries, follows links, and scrolls up and down web pages. It is trained to cite its sources, which makes it easier to give feedback to improve factual accuracy.

J. Hoffmann et al. (DeepMind) – "*Training Compute-Optimal Large Language Models*", https://arxiv.org/abs/2203.15556, March 2022

> We investigate the optimal model size and number of tokens for training a transformer language model under a given compute budget. We find that current large language models are significantly undertrained, a consequence of the recent focus on scaling language models whilst keeping the amount of training data constant.

J. Wei et al. (Google) – "*Language Models Perform Reasoning via Chain of Thought*", https://ai.googleblog.com/2022/05/language-models-perform-reasoning-via.html, May 2022

OpenAI Blog Entry – "*Aligning Language Models to Follow Instructions*", https://openai.com/blog/instruction-following/, January 2022

> We've trained language models that are much better at following user intentions than GPT-3 while also making them more truthful and less toxic, using techniques developed through our alignment research. These InstructGPT models, which are trained with humans in the loop, are now deployed as the default language models on our API.

J. Wei et al. (Google) – "*Chain of Thought Prompting Elicits Reasoning in Large Language Models*", https://arxiv.org/abs/2201.11903, January 2022, revised October 2022

> We explore how generating a chain of thought -- a series of intermediate reasoning steps -- significantly improves the ability of large language models to perform complex reasoning. In particular, we show how such reasoning abilities emerge naturally in sufficiently large language models via a simple method called chain of thought prompting, where a few chain of thought demonstrations are provided as exemplars in prompting.

OpenAI Blog Entry – "ChatGPT: Optimizing Language Models for Dialogue", https://openai.com/blog/chatgpt/, 30. November 2022

> We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer follow-up questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. ChatGPT is a sibling model to InstructGPT …

Y. Fu et al. (Univ. of Edinburgh) – "*How does GPT Obtain its Ability? Tracing Emergent Abilities of Language Models to their Sources*", How does GPT Obtain its Ability? Tracing Emergent Abilities of Language Models to their Sources (notion.site), December 2022

Wikipedia articles: GPT-3 – Wikipedia, ChatGPT – Wikipedia

References

Gwern – "*GPT-3 Creative Fiction*", https://www.gwern.net/GPT-3, 2020-06-19 – 2022-02-10

A. Thompson (Life Architect) – "*Inside language models (from GPT-3 to PaLM)*", https://lifearchitect.ai/models/, last updated 13. December 2022

A. Thompson (Life Architect) – "*Books by AI (GPT-3)*", https://lifearchitect.ai/books-by-ai/, last updated 14. December 2022

A. Thompson (Life Architect) – "*Use cases for large language models like GPT-3*", https://lifearchitect.ai/use-cases/, July 2022

A. Thompson (Life Architect) – " *What's in my AI?*", https://lifearchitect.ai/whats-in-my-ai/, March 2022, last updated 18. October 2022
> A Comprehensive Analysis of Datasets Used to Train GPT-1, GPT-2, GPT-3, GPT-NeoX-20B, Megatron-11B, MT-NLG, and Gopher

H.Hotz – "*I Used ChatGPT to Create an Entire AI Application on AWS*, https://towardsdatascience.com/i-used-chatgpt-to-create-an-entire-ai-application-on-aws-5b90e34c3d50, 02. December 2022
> … in this blog post I describe how I used ChatGPT to create a simple sentiment analysis application from scratch. The app should run on an EC2 instance and utilise a state-of-the-art NLP model from the Hugging Face Model Hub. The results were astonishing 😮

J. Degrave (DeepMind) – "*Building A Virtual Machine inside ChatGPT*", https://www.engraved.blog/building-a-virtual-machine-inside/, 03. Dec. 2022

J. Radoff – "*Creating a Text Adventure Game with ChatGPT*", https://medium.com/building-the-metaverse/creating-a-text-adventure-game-with-chatg-cffeff4d7cfd, 04. Dec. 2022

J. Wei et al. (Google Research, DeepMind, Stanford U., UNC Chapel Hill) – "*Emergent Abilities of Large Language Models*", https://arxiv.org/abs/2206.07682, October 2022

K. Steven – "*15 ChatGPT Examples (how to use)*" , https://khrisdigital.com/chatgpt-examples/, 12. Dec. 2022

Intercom – "*How ChatGPT changed everything: Thoughts from the frontline of the AI/ML revolution*", https://youtu.be/SCsSpqZq_xA
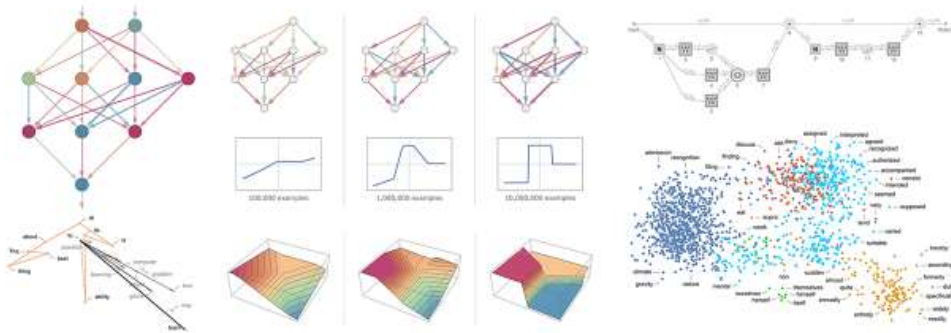
sentdex – "*OpenAI's ChatGPT is a MASSIVE step forward in Generative AI*", https://youtu.be/HTWfA7KFzoA

Yannic Kilcher – "*ChatGPT: This AI has a JAILBREAK?! (Unbelievable AI Progress)*", https://youtu.be/0A8ljAkdFtg

References: ChatGPT, GitHub Copilot X, Code Interpreter Plug-in

What Is ChatGPT Doing ... and Why Does It Work?

February 14, 2023

It's Just Adding One Word at a Time

→ What Is ChatGPT Doing … and Why Does It Work?—Stephen Wolfram Writings

(65) REAL Uses of ChatGPT As A Developer | 12 Practical Examples - YouTube

(65) GitHub Copilot X Explained | A big step forward... - YouTube

(65) GPT 4 Got Upgraded - Code Interpreter (ft. Image Editing, MP4s, 3D Plots, Data Analytics and more!) - YouTube

(65) 12 New Code Interpreter Uses (Image to 3D, Book Scans, Multiple Datasets, Error Analysis …) - YouTube

→ ChatGPT + Code Interpreter = Magic – @AndrewMayne (wordpress.com) – dated 23. March 2023

x

GPT-4 References

[Submitted on 15 Mar 2023 (v1), last revised 27 Mar 2023 (this version, v3)]

GPT-4 Technical Report

OpenAI

Download PDF

We report the development of GPT-4, a large-scale, multimodal model which can accept image and text inputs and produce text outputs. While less capable than humans in many real-world scenarios, GPT-4 exhibits human-level performance on various professional and academic benchmarks, including passing a simulated bar exam with a score around the top 10% of test takers. GPT-4 is a Transformer-based model pre-trained to predict the next token in a document. The post-training alignment process results in improved performance on measures of factuality and adherence to desired behavior. A core component of this project was developing infrastructure and optimization methods that behave predictably across a wide range of scales. This allowed us to accurately predict some aspects of GPT-4's performance based on models trained with no more than 1/1,000th the compute of GPT-4.

[2303.08774v3] GPT-4 Technical Report (arxiv.org)

(65) Sparks of AGI: early experiments with GPT-4 - YouTube

[Submitted on 22 Mar 2023 (v1), last revised 13 Apr 2023 (this version, v5)]

Sparks of Artificial General Intelligence: Early experiments with GPT-4

Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, Harsha Nori, Hamid Palangi, Marco Tulio Ribeiro, Yi Zhang

Download PDF

Artificial intelligence (AI) researchers have been developing and refining large language models (LLMs) that exhibit remarkable capabilities across a variety of domains and tasks, challenging our understanding of learning and cognition. The latest model developed by OpenAI, GPT-4, was trained using an unprecedented scale of compute and data. In this paper, we report on our investigation of an early version of GPT-4, when it was still in active development by OpenAI. We contend that (this early version of) GPT-4 is part of a new cohort of LLMs (along with ChatGPT and Google's PaLM for example) that exhibit more general intelligence than previous AI models. We discuss the rising capabilities and implications of these models. We demonstrate that, beyond its mastery of language, GPT-4 can solve novel and difficult tasks that span mathematics, coding, vision, medicine, law, psychology and more, without needing any special prompting. Moreover, in all of these tasks, GPT-4's performance is strikingly close to human-level performance, and often vastly surpasses prior models such as ChatGPT. Given the breadth and depth of GPT-4's capabilities, we believe that it could reasonably be viewed as an early (yet still incomplete) version of an artificial general intelligence (AGI) system. In our exploration of GPT-4, we put special emphasis on discovering its limitations, and we discuss the challenges ahead for advancing towards deeper and more comprehensive versions of AGI, including the possible need for pursuing a new paradigm that moves beyond next-word prediction. We conclude with reflections on societal influences of the recent technological leap and future research directions.

[2303.12712] Sparks of Artificial General Intelligence: Early experiments with GPT-4 (arxiv.org)

References – Healthcare Related
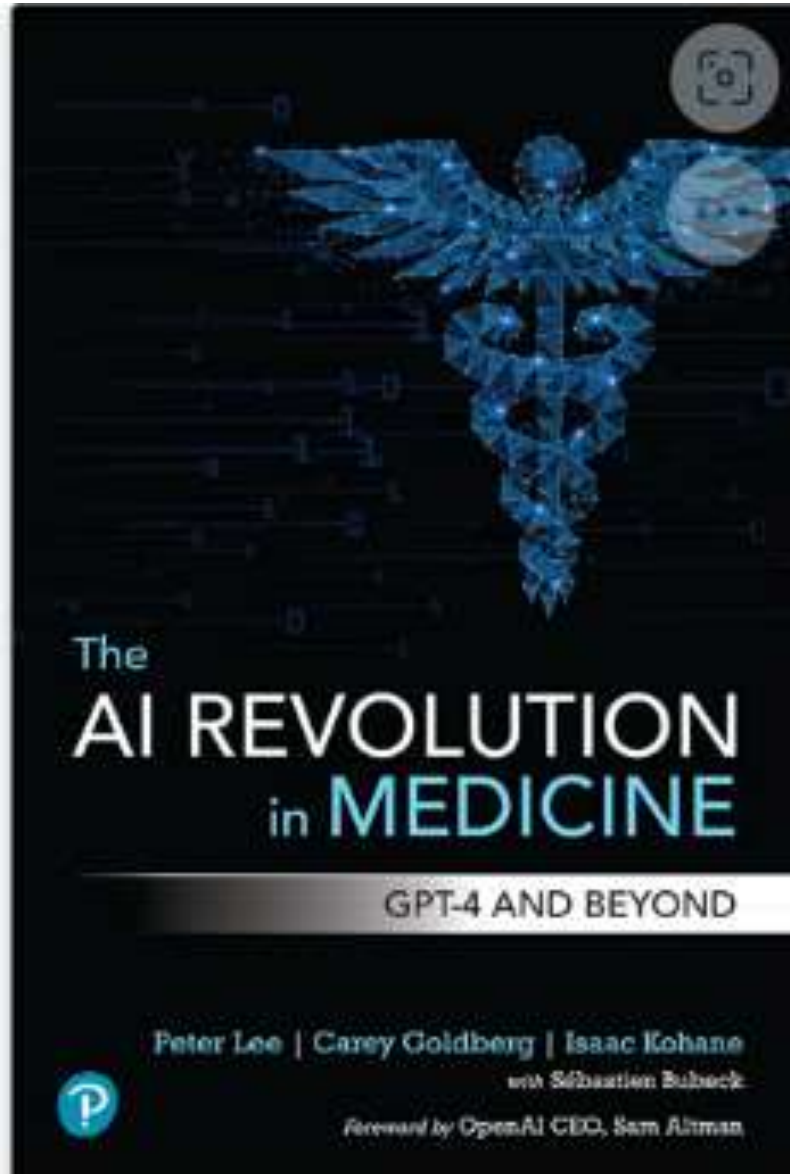
Peter Lee, Carey Goldberg, Isaac Kohane:
The AI Revolution in Medicine – GPT-4 and beyond

From Eric Topol's review ([The GPT-x Revolution in Medicine - by Eric Topol (substack.com)](#)):
The authors had 6 months to test drive GPT-4 before its release, specifically to consider its medical use cases, and put together their thoughts and experience to stimulate an important conversation in the medical community about the impact of AI. …

Before getting into the book's content, let me comment on its authors. I've known Peter Lee, a very accomplished computer scientist who heads up Microsoft Research, for many years. He's mild-mannered, not one to get into hyperbole. Yes, Microsoft has invested at least $10 billion for a major stake in OpenAI, so there's no doubt of a perceived conflict, but that has not affected my take from the book. He brought on Zak Kohane, a Harvard pediatric endocrinologist and data scientist, as a co-author. Zak is one of the most highly regarded medical informatic experts in academic medicine and is the editor-in-chief of the new NEJM-AI journal. As with Peter, I've gotten to know Zak over many years, and he is not one to exaggerate—he's a straight shooter. So the quote at the top of this post coming from him is significant. The third author is Carey Goldberg, a leading health and science journalist who previously was the Boston bureau chief for the New York Times. …

The start of the book is a futuristic vision of medical chatbots. A second-year medical resident who, in the midst of a crashing patient, turns to the GPT-4 app on her smartphone for help. While it's all too common these days on medical rounds for students and residents to do Google searches, this is a whole new look. GPT-4 analyzes the patient's data and provides guidance for management with citations. Beyond this patient, GPT-4 helps the resident for pre-authorizations, writes a discharge summary to edit and approve, provides recommendations for a clinical trial for one of her patients, reviews the care plans for all her patients, and provides feedback on her own health data for self-care.
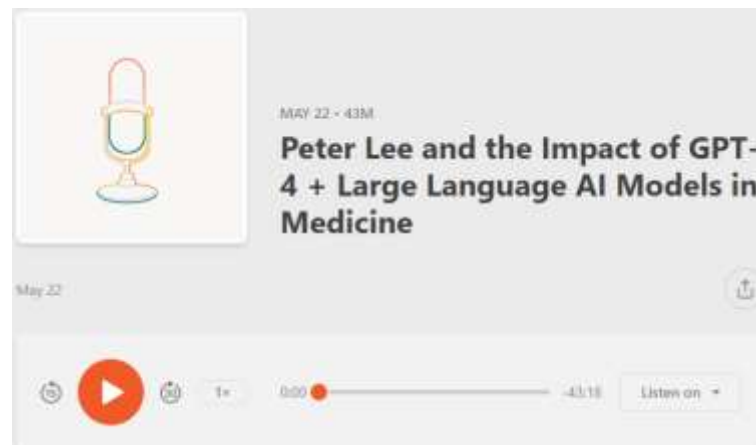
The AI Revolution in Medicine: GPT-4 and Beyond: Lee, Peter, Goldberg, Carey, Kohane, Isaac: 9780138200138: Amazon.com: Books

References Related to Healthcare

[(65) Allen School Distinguished Lecture: Peter Lee (Microsoft Research & Incubations) - YouTube](#)

[(65) Dr Peter Lee (Microsoft VP Research) • The Emergence of General AI for Medicine. - YouTube](#)

[Peter Lee and the Impact of GPT-4 + Large Language AI Models in Medicine (substack.com)](#)

[Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine | NEJM](#)
Requires subscriptions to the New England Journal of Medicine

[The GPT-4 Episode: Microsoft's Peter Lee on the Future of Language Models in Medicine | NEJM AI Grand Rounds](#)
Podcast episode from 29 March 2023