Week 15: Generative AI
April 19, 2023
Introducing ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.
ANNALS OF TECHNOLOGY

CHATGPT IS A BLURRY JPEG OF THE WEB

OpenAI’s chatbot offers paraphrases, whereas Google offers quotes. Which do we prefer?

By Ted Chiang
February 9, 2023
Some Glimpse AGI in ChatGPT. Others Call It a Mirage

A new generation of AI algorithms can *feel* like they’re reaching artificial general intelligence—but it’s not clear how to measure that.
People Are Using AI for Therapy, Even Though ChatGPT Wasn’t Built for It

Some users see it as a way to supplement traditional mental health services, despite troubling privacy implications.
Foundation models for generalist medical artificial intelligence

Michael Moor, Oishi Banerjee, Zahra Shakeri Hossein Abad, Harlan M. Krumholz, Jure Leskovec, Eric J. Topol & Pranav Rajpurkar

Nature 616, 259–265 (2023) | Cite this article

Abstract

The exceptionally rapid development of highly flexible, reusable artificial intelligence (AI) models is likely to usher in newfound capabilities in medicine. We propose a new paradigm for medical AI, which we refer to as generalist medical AI (GMAI). GMAI models will be capable of carrying out a diverse set of tasks using very little or no task-specific labelled data. Built
ChatGPT may be coming for our jobs. Here are the 10 roles that AI is most likely to replace.

Aaron Mok and Jacob Zinkula  Apr 9, 2023, 11:36 AM
ChatGPT passes MBA exam given by a Wharton professor

The bot’s performance on the test has “important implications for business school education,” wrote Christian Terwiesch, a professor at the University of Pennsylvania’s Wharton School.
Synthetic Lies: Understanding AI-Generated Misinformation and Evaluating Algorithmic and Human Solutions

Jiawei Zhou  
Georgia Institute of Technology  
Atlanta, GA, USA  
j.zhou@gatech.edu

Yixuan Zhang  
Georgia Institute of Technology  
Atlanta, GA, USA  
yixuan@gatech.edu

Qianni Luo  
Ohio University  
Athens, OH, USA  
ql047311@ohio.edu

Andrea G Parker  
Georgia Institute of Technology  
Atlanta, GA, USA  
andrea@cc.gatech.edu

Munmun De Choudhury  
Georgia Institute of Technology  
Atlanta, GA, USA  
munmund@gatech.edu

ABSTRACT
Large language models have abilities in creating high-volume human-like texts and can be used to generate persuasive misinformation. However, the risks remain under-explored. To address the gap, this work first examined characteristics of AI-generated misinformation (AI-misinfo) compared with human creations, and then evaluated the applicability of existing solutions. We compiled human-created COVID-19 misinformation and abstracted it into narrative prompts for a language model to output AI-misinfo. We found significant linguistic differences within human-AI pairs, and patterns of AI-misinfo in enhancing details, communicating uncertainties, drawing conclusions, and simulating personal tones. While existing models remained capable of classifying AI-misinfo, a significant performance drop compared to human-misinfo was observed. Re-

1 INTRODUCTION
The Coronavirus Disease (COVID-19) pandemic has brought attention to the proliferation of health misinformation\(^1\). From fake cures to conspiracy theories, misinformation has led to substantial adverse effects at the individual as well as societal levels. Examples of such effects include mortality and hospital admissions [20, 48], public fear and anxiety [79, 107], eroded trust in health institutions [87], and exacerbated racial discrimination and stigma [41, 48]. Finding ways to combat misinformation is therefore of critical importance from the perspectives of both public health and governance. Manual identification of misinformation is, however, extremely laborious and often does not scale: a key issue given the rise of misinformation on social media [71]. As such, artificial intelligence (AI) techniques have been touted as a timely and scalable solution for
ChatGPT banned in Italy over privacy concerns

1 April