

Impact of LLMs

Indira Sen, SILLM Lecture 10

Agenda

- Threshold Effects
- Types of LLM impact
 - Scientific
 - Societal
 - Education
 - Art
 - Case study 1: jobs
 - Case study 2: public goods
 - Environmental
 - ...
 - ‘Existential risks’

Recap: The different ways we've seen LLMs being used

CHATGPT OUTPERFORMS CROWD-WORKERS
FOR TEXT-ANNOTATION TASKS

Social Simulacra: Creating Populated Prototypes for Social Computing Systems

AI Psychometrics: Assessing the Psychological Profiles of Large Language Models Through Psychometric Inventories

Perspectives on Psychological Science

1–19

© The Author(s) 2023



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/17456916231214460
www.psychologicalscience.org/PPS

Using Large Language Models to Simulate Multiple Humans and Replicate Human Subject Studies

Gati Aher¹ Rosa I. Arriaga² Adam Tauman Kalai³

Abstract

We introduce a new type of text called “Training

trolled experiments, and we thus avoid philosophical debates around the meaning of “understanding” (Bender &

M

ABSTRACT
Max
Beat
¹Business
²Department
³Department
⁴Compl

Social comp...
arise in an...
is currently...
nately, man...
at a larger...
might beha...
We ill...
sign before...
be rep...
duce social...
assum...

Abstr

Recap: other ways LLMs are being used

Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum

John W. Ayers, PhD,

» Author Affiliation

JAMA Intern Med. 2023

Grounded Copilot: How Programmers Interact with Code-Generating Models

SHRADDHA

MICHAEL

NADIA PC

Powered by
the face of p
theory analy
prior experie
finding is th
knows what
how to proce
for improvini

Jürgen Rudolph^A

A

Director of Research, Kaplan Singapore

Shannon Tan^B

B

Research Executive, Kaplan Singapore

Samson Tan^C

C

Director of Regional Strategy & Operations (Singapore), Civica Asia Pacific

DOI: <https://doi.org/10.37074/jalt.2023.6.1.23>

Key Points

Question Can an
comparable qualit

Findings In this c
team of licensed
CCS Concept
engineering
Additional K

Abstract

Developments in the chatbot space have been accelerating at breakneck speed since late November 2022. Every day,

launch in November 2022 (Rudolph et al., 2023). As recent faddish exuberances around blockchain, cryptos, initial coin offerings, the metaverse, and non-fungible tokens have

Others?

Threshold Effects

<https://crfm.stanford.edu/workshop.html>

<https://www.youtube.com/watch?v=lx1MExMIAk&t=2056s>

The image shows a YouTube video thumbnail. The main title "A STORY OF ANCIENT HISTORY" is at the top, followed by the specific video title "Threshold Effects". Below the titles, it says "Michael Bernstein Stanford Computer Science". On the right side, there is a video frame showing a man (Michael Bernstein) speaking outdoors with a sunset background. The thumbnail includes the HAI logo (Stanford University Center for Research on Foundation Models), the speaker's name and title, and the event details: "Workshop on Foundation Models | August 23-24, 2021 | #foundationmodels". At the bottom, there are standard YouTube interaction buttons for likes, shares, and downloads.

Types of LLM Impact

- Scientific
- Societal
 - Education
 - Case study 1: art
 - Case study 2: jobs
- Environmental
- ...
- ‘Existential risks’

Types of LLM Impact: Scientific

'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures

Google's deep-learning program for determining the 3D shapes of proteins stands to transform biology, say scientists.

By [Ewen Callaway](#)



The world's most widely adopted AI developer tool.

Get started with Copilot >

Where knowledge begins

Ask anything...

New Thread Ctrl +

Q Home

Discover

Library

Sign In

Sign Up

Try Pro

Upgrade for image upload, smarter AI, and more Copilot.

Learn More

Focus Attach Copilot

Try asking The best fishing spots in Canada Alternatives to an alarm clock
Who invented the plunger? What causes a rainbow?

Try Pro Careers Playground Blog

Types of LLM Impact: Scientific

- LLMs can help scientific progress by improving efficiency
- Github CoPilot for Pair programming
- Summarize and simplify scientific literature

Types of LLM Impact: Education

Educators Battle Plagiarism As 89% Of Students Admit To Using OpenAI's ChatGPT For Homework

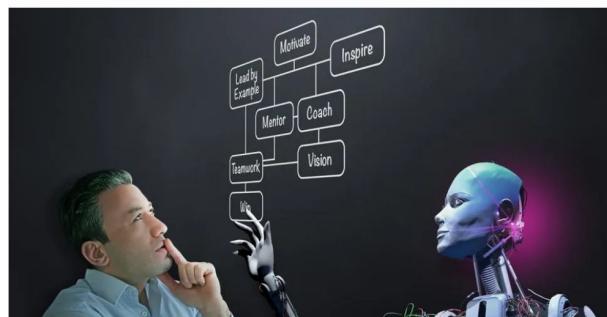
Chris Westfall Contributor 

Guidance for leaders and aspiring leaders, interested in career impact

Follow



Jan 28, 2023, 01:37pm EST



Types of LLM Impact: Education

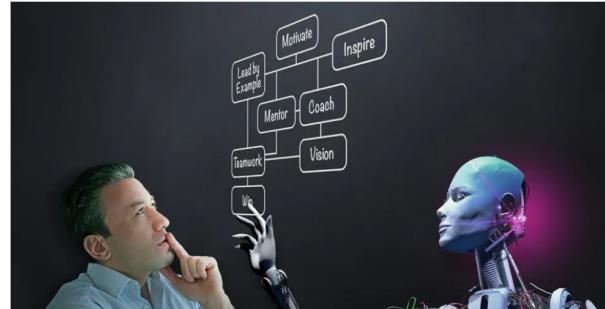
Educators Battle Plagiarism As 89% Of Students Admit To Using OpenAI's ChatGPT For Homework

Chris Westfall Contributor 

Guidance for leaders and aspiring leaders, interested in career impact

Follow

Jan 28, 2023, 01:37pm EST



BOT BUST

Professor Flunks All His Students After ChatGPT Falsey Claims It Wrote Their Papers

Texas A&M University–Commerce seniors who have already graduated were denied their diplomas because of an instructor who incorrectly used AI software to detect cheating

BY MILES KLEE

MAY 17, 2023



Types of LLM Impact: Education

- 33-42% students report using ChatGPT
- 50-60% teachers use it

[Teachers are embracing ChatGPT—even more than students](#)

Types of LLM Impact: Education

- 33-42% students report using ChatGPT
- 50-60% teachers use it
- AI-generated content detectors:
 - Hard to establish validity
 - Biased against non-native speakers [see [GPT detectors are biased against non-native English writers](#)]
 - Possible for specific use cases? [see [Artificial Artificial Artificial Intelligence](#)]

As of July 20, 2023, the AI classifier is no longer available due to its low rate of accuracy. We are working to incorporate feedback and are currently researching more effective provenance techniques for text, and have made a commitment to develop and deploy mechanisms that enable users to understand if audio or visual content is AI-generated.

[https://openai.com/blog/new-ai-classifier-for-i
ndicating-ai-written-text](https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text)

[Teachers are embracing ChatGPT—even more than students](#)

Types of LLM Impact: Art*

- Rapid evolution in AI art
- Impact on artists
- Ownership, credit, and copyright
- Also applies to books, music, screenplays, movies, ...



Trump vs.
Obama
playing
Basketball

Midjourney,
March 2022



Midjourney,
March 2023

Image credit: ainterestingaf on Instagram

Types of LLM Impact: Jobs

"AI made us do it" is Big Tech's new
layoff rationale



Megan Morrone

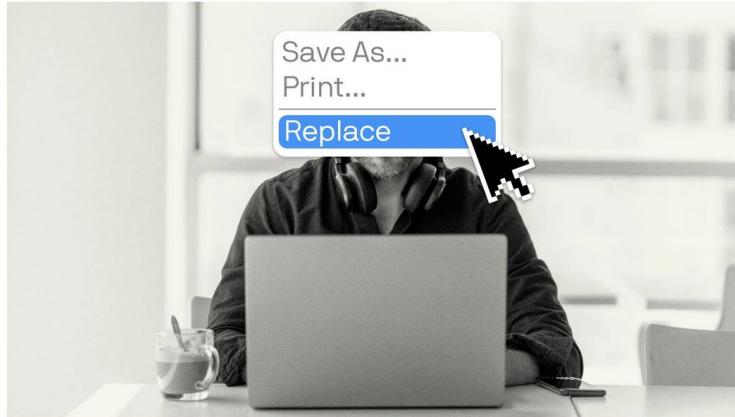


Illustration: Shoshana Gordon/Axios

Big tech companies are pointing to a renewed focus on AI as a justification for [layoffs](#).

Why it matters: The new trend in tech company layoffs could be less about replacing workers with AI, and more about replacing workers with a smaller number of workers who are [more skilled in AI](#) for now.

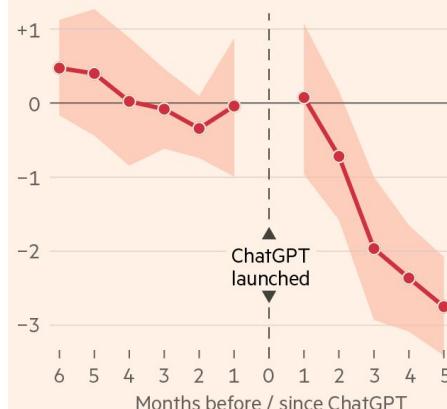
Types of LLM Impact: Jobs

- Impacts on Skilled Labor, including freelancers
- freelancers whose work was the most exposed to AI earned over 5% less in April 2023
- skill-biased technological change (SBTC) framework assumes ‘cognitive non-routine (high-skill) workers’ are not replaced, but augmented.
- But more skills != more safety; top freelancers are disproportionately affected by AI

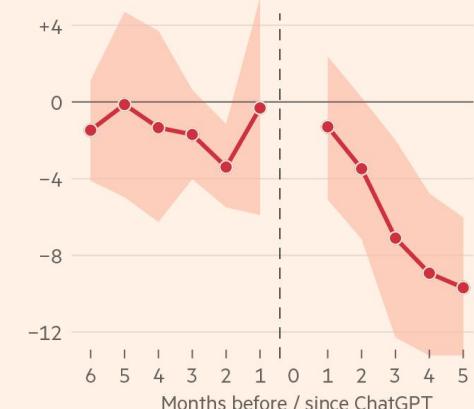
Generative AI is already taking white-collar jobs and wages in the online freelancing world

Change in employment and earnings from writing and editing jobs on an online freelancing platform after the launch of ChatGPT

% change in monthly freelance jobs ...



... and earnings



Source: *The Short-Term Effects of Generative AI on Employment: Evidence from an Online Labor Market* (Hui et al, 2023)
© FT

[The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market](#)

Types of LLM Impact: Jobs

Chatbot to Replace Human Staffers at National Eating Disorders Association Helpline

Volunteers say the move, which came after staffers voted to unionize, will harm callers to the helpline

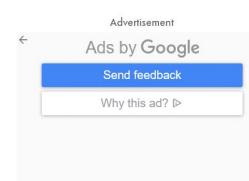
By Cara Lynn Shultz | Published on May 24, 2023 11:18AM EDT



National Eating Disorders Association helpline is replacing volunteers with a chatbot. PHOTO: GETTY

The National Eating Disorders Association (NEDA) is disbanding its helpline in its current form, replacing hundreds of staffers and volunteers with a chatbot named Tessa.

Nearly 70,000 people reached out to the helpline last year, according to a recent [NPR report](#). Those numbers doubled during the COVID-19 pandemic, when volunteers and staffers recalled more "crisis-type"



Types of LLM Impact: Jobs

Chatbot to Replace Human Staffers at National Eating Disorders Association Helpline

Volunteers say the move, which came after staffers voted to unionize, will harm callers to the helpline

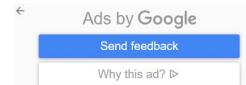
By Cara Lynn Shultz | Published on May 24, 2023 11:18AM EDT



National Eating Disorders Association helpline is replacing volunteers with a chatbot. PHOTO: GETTY

The National Eating Disorders Association (NEDA) is disbanding its helpline in its current form, replacing hundreds of staffers and volunteers with a chatbot named Tessa.

Nearly 70,000 people reached out to the helpline last year, according to a recent NPR report. Those numbers doubled during the COVID-19 pandemic, when volunteers and staffers recalled more "crisis-type"



US eating disorder helpline takes down AI chatbot over harmful advice

National Eating Disorder Association has also been under criticism for firing four employees in March who formed a union



Tessa, which Neda claims was never meant to replace the helpline workers, almost immediately ran into issues. Photograph: redsnapper/Alamy

The National Eating Disorder Association (Neda) has taken down an artificial intelligence chatbot, "Tessa", after reports that the chatbot was providing harmful advice.

Types of LLM Impact: Jobs

- NEDA: National Eating Disorders Association
- worked with psychology researchers and Cass AI, a company that develops AI chatbots focused on mental health, to develop ‘Tessa’
- Fired employees because they were unionizing and replaced them with Tessa
- Tessa gave them harmful weight loss advice; particularly bad for an eating disorder helpline

Types of LLM Impact: Public Goods

- **Public goods:** a kind of common good; two main properties:
 - cannot prevent someone from consuming the good.
 - If a person consumes the good, this does not reduce the availability of the good to others.
- Digital Public Goods: Stackoverflow (Stackexchange in general), advice-related subreddits (r/LPT)

Types of LLM Impact: Public Goods

- **Public goods:** a kind of common good; two main properties:
 - cannot prevent someone from consuming the good.
 - If a person consumes the good, this does not reduce the availability of the good to others.
- Digital Public Goods: **Stackoverflow** (Stackexchange in general), advice-related subreddits (r/LPT)

Are Large Language Models a Threat to Digital Public Goods? Evidence from Activity on Stack Overflow

Maria del Rio-Chanona ^{1,2}, Nadzeya Laurentsyeva ³, and Johannes Wachs ^{4,5,1,*}

¹ Complexity Science Hub, Vienna.

² Harvard Kennedy School

³ Faculty of Economics, LMU Munich

⁴ Corvinus University of Budapest

⁵ Centre for Economic and Regional Studies, Hungary *

2023-07-14

Abstract

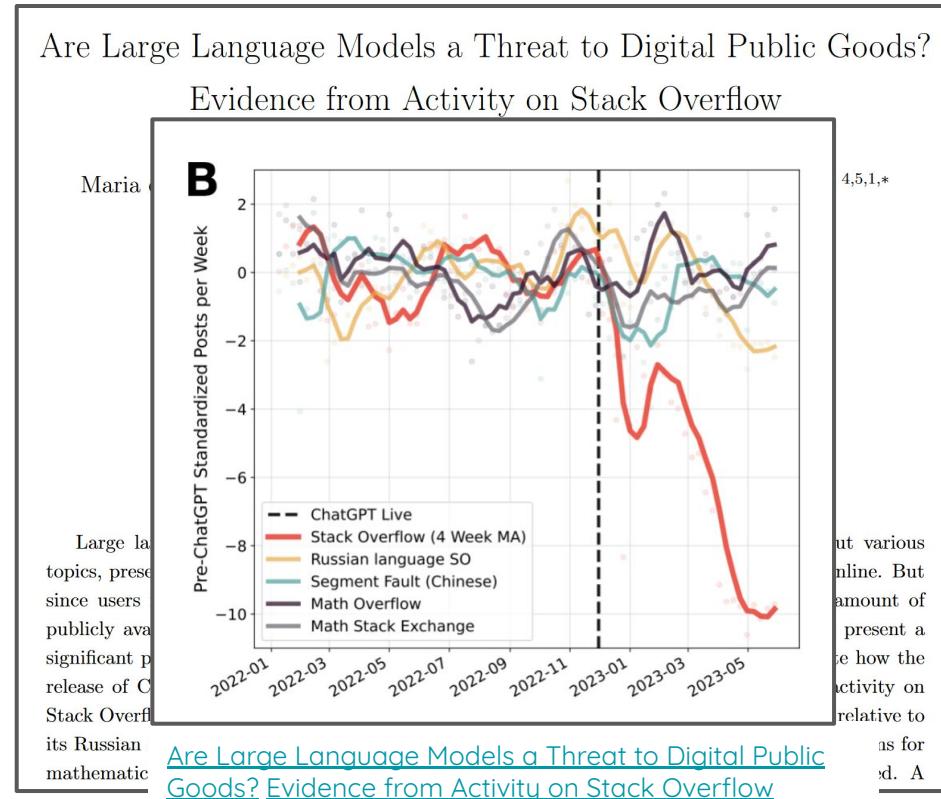
Large language models like ChatGPT efficiently provide users with information about various topics, presenting a potential substitute for searching the web and asking people for help online. But since users interact privately with the model, these models may drastically reduce the amount of publicly available human-generated data and knowledge resources. This substitution can present a significant problem in securing training data for future models. In this work, we investigate how the release of ChatGPT changed human-generated open data on the web by analyzing the activity on Stack Overflow, the leading online Q&A platform for computer programming. We find that relative to its Russian mathematical

[Are Large Language Models a Threat to Digital Public Goods? Evidence from Activity on Stack Overflow](#)

as for
d. A

Types of LLM Impact: Public Goods

- **Public goods:** a kind of common good; two main properties:
 - cannot prevent someone from consuming the good.
 - If a person consumes the good, this does not reduce the availability of the good to others.
- Digital Public Goods: **Stackoverflow** (Stackexchange in general), advice-related subreddits (r/LPT)
- traffic to Stackoverflow reduced after ChatGPT's introduction
- **Consequence:** reduction in open data for people and models



Types of LLM Impact: Environmental Effects

- Training LLMs have a carbon footprint
=> training one BERT model == one trans-atlantic flight [[How to Make Generative AI Greener](#)]
- Also, if you train on the cloud; affected by the region of your cluster
- **Energy-efficient LLMs:**
 - BLOOM [see [Lucconi et al.](#).]
 - “Training LLaMA 13B emits 24 times less greenhouse gases than training GPT-3 175B yet performs better on benchmarks.” [[Yann leCun](#)]

AI'S CARBON FOOTPRINT

The emissions associated with training the language-learning model BERT depend on the time of year, and on the location of the data centre.

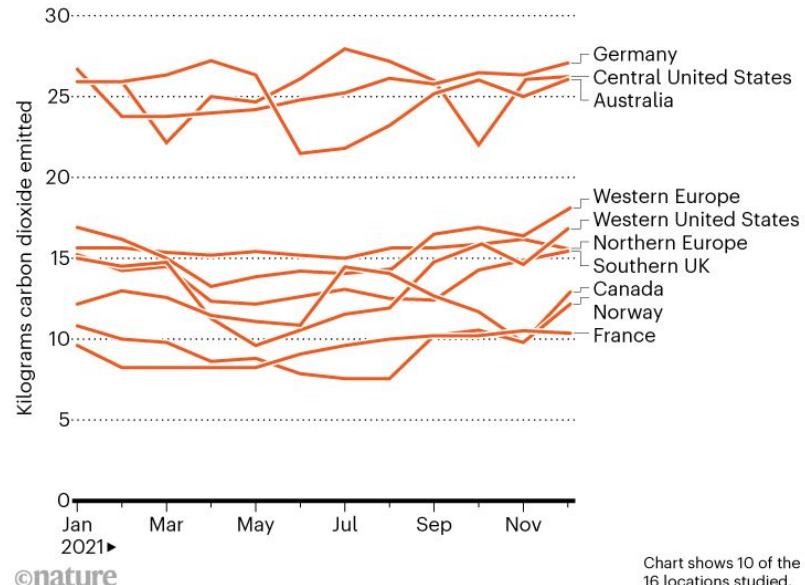


Chart shows 10 of the 16 locations studied.
©nature

[How to shrink AI's ballooning carbon footprint](#)

Existential Risks of LLMs



Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

[View this open letter online.](#)

Published

March 22, 2023

PDF created

May 5, 2023

Signatures

27565

AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research¹ and acknowledged by top AI labs.² As stated in the widely-endorsed [Asilomar AI Principles](#), Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources.

Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control.

Should we let machines flood our information channels with propaganda and untruth?

Should we automate away all the jobs, including the fulfilling ones?

Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization?

Existential Risks of LLMs



Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

[View this open letter online.](#)

Published

March 22, 2023

PDF created

May 5, 2023

Signatures

27565

AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research¹ and acknowledged by top AI labs.² As stated in the widely-endorsed [Asilomar AI Principles](#), Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources.

Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control.

Should we let machines flood our information channels with propaganda and untruth?

Should we automate away all the jobs, including the fulfilling ones?

Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization?

Summary

- LLMs have already started having impacts on us personally and societally
- LLMs offer *low(er) thresholds* and *high ceilings*:
 - They make AI easily accessible (though some education required for prompting)
 - They have many capabilities
 - High chances of impactful co-option
- clear positive and negative impacts on politics, education, labor, public goods, art, the environment and many more...
- We do not see clear ways of reducing negative impacts, *yet*

Summary

- LLMs have already started having impacts on us personally and societally
- LLMs offer *low(er) thresholds and high ceilings*:
 - They make AI easily accessible (though some education required for prompting)
 - They have many capabilities
 - High chances of impactful co-option
- clear positive and negative impacts on politics, education, labor, public goods, art, the environment and many more...
- We do not see clear ways of reducing negative impacts, yet

FORBES > INNOVATION > ENTERPRISE TECH

The Hot New Job That Pays Six Figures: AI Prompt Engineering

Bernard Marr Contributor 



May 11, 2023, 02:25am EDT



Back in 2017, a [report](#) by Dell Technologies and the Institute Of The Future stated that 85% of the jobs that will exist in 2030 haven't been invented yet.



Further Reading

- On the opportunities and risks of foundation models [See Section 5]
- What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature
- Synthetic Lies: Understanding AI-Generated Misinformation and Evaluating Algorithmic and Human Solutions
- Energy and policy considerations for deep learning in NLP
- The Alt-Right Manipulated My Comic. Then A.I. Claimed It. [AI Art and its Impact on artists]
- <https://www.aisnakeoil.com/p/a-misleading-open-letter-about-sci> [real risks vs. speculative risks]

Next Lecture: Critique of LLMs

Readings:

1. [On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?](#)
1. [ChatGPT Is a Blurry JPEG of the Web](#) (email us if you can't access this)

Questions for discussion:

- What can LLMs do well? What can't they?
- How should LLMs be regulated?
- How to avoid and mitigate adverse effects of LLMs in our society?