

MPRA

Munich Personal RePEc Archive

Long Tails and the Impact of GPT on Labor

Kausik, B.N.

29 March 2023

Online at <https://mpra.ub.uni-muenchen.de/117135/>
MPRA Paper No. 117135, posted 26 Apr 2023 09:19 UTC

Long Tails & the Impact of GPT on Labor

B.N. Kausik¹

Mar 29 2023

Abstract

Recent advances in AI technologies renew urgency to the question whether automation will cause mass unemployment and reduction in standards of living. While prior work analyzes historical economic data for the impact of automation on labor, we seek a test to predict the impact of emerging automation technologies such as Generative Pretrained Transformers (GPT). Towards that goal, we observe that human needs favor long tail distributions, i.e., a long list of niche items that are substantial in aggregate popularity. In turn, the long tails are reflected in the products and services that fulfill those needs. Technologies that address a small portion of the distribution, typically the head, free up human labor to focus on more complex tasks in the long tail, thereby improving productivity and potentially lifting wages. In contrast, technologies that cover substantial portions of the long tail can squeeze wages or displace humans entirely. With this in mind, we propose a long tail test for automation technologies to predict their impact on labor. We find that popular GPTs perform poorly on such tests in that they are erratic on straightforward long tail tasks, hence absent breakthroughs, will augment human productivity rather than cause mass displacement of human labor. Going forward, we believe that to have a broad impact on displacing or devaluing human labor, AI must at least be capable of long-tail tasks that humans perform with ease.

JEL Classification: J21, J33, O33

¹ Conflict disclosure: unaffiliated independent.

Author's bio: <https://www.linkedin.com/in/bnkausik/> Contact: bnkausik@gmail.com

Introduction

Recent advances in AI technologies such as Large Language Models (LLMs) or Generative Pre-trained Transformers (GPTs), e.g., Google's Bard and OpenAI's ChatGPT, bring new urgency to the question whether automation will cause mass unemployment and reduction in standards of living. At first glance, these AI technologies display amazing human-like intelligence, although with flashes of absurd “hallucinations,” Fowler, (2023) and Roose, (2023). We consider the question whether these technologies are at the point where human labor is displaced or devalued, rather than augmented.

Korinek and Stiglitz (2020) write:

Rapid progress in new technologies such as Artificial Intelligence has recently led to widespread anxiety about potential job losses. This paper asks how to guide innovative efforts so as to increase labor demand and create better-paying jobs. We develop a theoretical framework to identify the properties that make an innovation desirable from the perspective of workers, including its technological complementarity to labor, the factor share of labor in producing the goods involved, and the relative income of the affected workers. Examples of labor-friendly innovations are intelligent assistants who enhance the productivity of human workers. The paper also discusses measures to steer technological progress in a desirable direction for workers, ranging from nudges for entrepreneurs to changes in tax, labor market and intellectual property policies to direct subsidies and taxes on innovation. In the future, we find that progress should increasingly be steered to provide workers with utility from the non-monetary aspects of their jobs.

Brynjolfsson (2022) writes:

In 1950, Alan Turing proposed an “imitation game” as the ultimate test of whether a machine was intelligent: could a machine imitate a human so well that its answers to questions are indistinguishable from those of a human. Ever since, creating intelligence that matches human intelligence has implicitly or explicitly been the goal of thousands of researchers, engineers and entrepreneurs. The benefits of human-like artificial intelligence (HLAI) include soaring productivity, increased leisure, and perhaps most profoundly, a better understanding of our own minds. But not all types of AI are human-like—in fact, many of the most powerful systems are very different from humans—and an excessive focus on developing and deploying HLAIs can lead us into a trap. As machines become better substitutes for human labor, workers lose economic and political bargaining power and become increasingly dependent on those who control the technology. In contrast, when AI is focused on augmenting humans rather than mimicking them, then humans retain the power to insist on a share of the value created. What's more, augmentation creates new capabilities and new products and services, ultimately generating far more value than merely human-like AI. While both types of AI can be enormously beneficial, there are currently excess incentives for automation rather than augmentation among technologists, business executives, and policymakers.

And Acemoglu (2022) writes:

This essay discusses several potential economic, political and social costs of the current path of AI technologies. I argue that if AI continues to be deployed along its current trajectory and remains unregulated, it may produce various social, economic and political harms. These include: damaging competition, consumer privacy and consumer choice; excessively automating work, fueling inequality, inefficiently pushing down wages, and failing to improve worker productivity; and damaging political discourse, democracy's most fundamental lifeblood.

More recently and in the context of GPT, Elondou et al (2023) write:

Our findings reveal that around 80% of the U.S. workforce could have at least 10% of their work tasks affected by the introduction of LLMs, while approximately 19% of workers may see at least 50% of their tasks impacted. We do not make predictions about the development or adoption timeline of such LLMs. The projected effects span all wage levels, with higher-income jobs potentially facing greater exposure to LLM capabilities and LLM-powered software. Significantly, these impacts are not restricted to industries with higher recent productivity growth. Our analysis suggests that, with access to an LLM, about 15% of all worker tasks in the US could be completed significantly faster at the same level of quality. When incorporating software and tooling built on top of LLMs, this share increases to between 47 and 56% of all tasks.

On a more optimistic note, Autor et al (2020) write:

No compelling historical or contemporary evidence suggests that technological advances are driving us toward a jobless future. On the contrary, we anticipate that in the next two decades, industrialized countries will have more job openings than workers to fill them, and that robotics and automation will play an increasingly crucial role in closing these gaps. Nevertheless, the impact of robotics and automation on workers will not be benign. These technologies, in concert with economic incentives, policy choices, and institutional forces, will alter the set of jobs available and the skills they demand.

And Basu, (2022) presents historical reasons for optimism:

Major technological advances cause turbulence, but if we manage to navigate the turbulence and come out of it, as humanity has done in the past, we end up growing faster rather than slower. Table 1 is a reminder of this in the context of the Industrial Revolution. Maddison's (2007) estimates suggest that, for two to three hundred years preceding the Industrial Revolution, the world was used to an average annual growth rate of 0.32 percent. This was more or less so until the early decades of the Industrial Revolution, before the benefits of better technology spread widely. After the dust and chaos of the Industrial Revolution settled, the new laws and regulations were in place, and a dramatic change in our economic life followed this technological advance spanning roughly one century from the mid-18th century to the mid-19th century. This is captured by the second column of Table 1. Western Europe transformed from a virtually stagnant economy to one bounding ahead at a growth rate of 2.11 percent. From 1870 to the eve of World War I, the

global economy had an annual growth rate of 2.12 percent. Growth accelerated further in the post-war period. In short, post-Industrial-Revolution global economic growth was more than eight times the growth that was normal before the revolution.

GDP growth rates (percent per annum)			
Country/Region	1500-1820	1870-1913	1973-2003
Western Europe	0.40	2.11	2.19
United States	0.86	3.94	2.94
China	0.41	-0.37	7.34
India	0.19	0.38	5.20
Africa	0.15	1.32	2.97
World	0.32	2.12	3.17

Table 1: Growth takeoff after the Industrial Revolution
(source: Maddison (2007) via Basu (2022))

In this paper, we ask whether we should be optimistic like Autor et al (2020) and Basu (2022) or pessimistic like the others. In particular, will GPTs devalue human labor causing mass unemployment and underemployment? Or, are they just another step in augmenting human productivity? Prior work, e.g., Acemoglu and Restruopo (2019), Autor et al (2022), analyzes historical economic data to study the impact of automation on labor. In contrast, we seek a test to predict the impact of emerging automation technologies such as GPT.

Our inquiry is structured as follows. First, we observe that human needs favor long tail distributions, demanding the same from the products and services that fulfill those needs. Secondly, we note that technologies that address a small portion of the distribution, typically the head, free up human labor to focus on more complex tasks in the long tail, thereby improving productivity and potentially lifting wages. For example, despite the large-scale adoption of personal computers and accounting software, the number of accountants and their real wages increased in recent decades. In contrast, technologies that cover substantial portions of the long tail can squeeze wages or displace humans entirely. For example, automated telephone exchanges completely replaced human telephone operators, since they addressed the problem across its full distribution. With this in mind, we propose a long tail test for automation technologies to predict their impact on labor. We find that popular GPTs perform poorly on such tests in that they are erratic on straightforward long tail tasks, hence absent breakthroughs, will augment human productivity rather than cause mass displacement of human labor.

Going forward, we believe that to have a broad impact on displacing or devaluing human labor, AI must at least be capable of long-tail tasks that humans perform with ease.

The Long Tail of Human Needs

The conventional wisdom of the so-called 80/20 rule, also known as the Pareto principle, is that 80% of the results accrue from 20% of the causes. The 20% is the “head” of the distribution, and the remaining 80% of the distribution is the “long tail” with a weak return on effort and not worth pursuing.

However, when it comes to human needs and tastes, the Pareto principle seems at best transient. For example, during the early years of the mass market automobile, the Ford Model T enjoyed a market share of almost 50%. However, as the market expanded and matured, competitors emerged to cater to the fragmented diversity of human tastes and needs. As of 2022, the top 20% of automobile brands represent ~60% of vehicles sold in the US market, see Fig. 1.

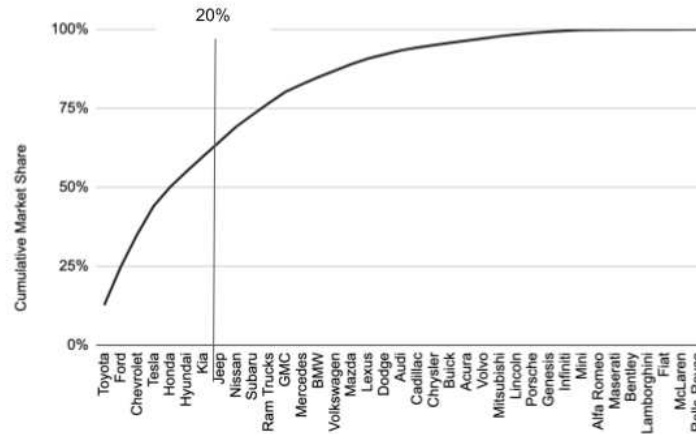


Fig. 1: The long tail of 2022 US vehicle sales.

With the advent of the internet, it became economically feasible to extend the long tail in many areas of human interest. Anderson (2004) recognized the importance of the long tail, and thereafter Anderson (2006) and Brynjolfsson et al (2007) expanded on the ideas in the context of retail. Goel et al (2010) present an analysis of the long-tail in retail, music, movies and web search, exploring whether most users' needs are met by popular items, i.e. the head of the distribution, or whether most users need popular items plus some eccentric items in the long tail of the distribution:

The first, and generally accepted theory, is that a majority of consumers prefer popular offerings while only a minority seek niche content; the second hypothesis is that everyone is a bit eccentric, consuming both popular and specialty products. These two theories, importantly, predict substantively different tradeoffs between inventory size and user satisfaction. In the former case, a small inventory of popular items would satisfy most people nearly all of the time, while in the latter, such an inventory would frustrate most people at least some of the time.

To distinguish between these possible alternatives, we examine extensive data on user preferences for movies, music, Web search, and Web browsing. In all of these domains, we find overwhelming evidence that nearly everyone is at least a bit eccentric. Our findings suggest an additional factor for the success of infinite-inventory retailers, that is, tail availability may boost head sales by

offering consumers the convenience of “one-stop shopping” for both their mainstream and niche interests.

In brief, Goel et al (2020) explain the success of online retailers and search engines that cover the long tail, over others that focus on the head and ignore the long tail. In Fig. 2, adapted from Goel et al (2010), the dotted and solid curves show the percentage of users whose needs are met 90% and 100% of the time respectively, as a function of the cumulative popularity of the items in the inventory ranked in descending order of popularity. The steep curves suggest that small increases in covering the long tail can result in disproportionately large increases in user satisfaction. Referring to Fig. 2, it is clear that at ~20% cumulative popularity, only a negligible percentage of users have much of their needs satisfied, violating the Pareto principle.

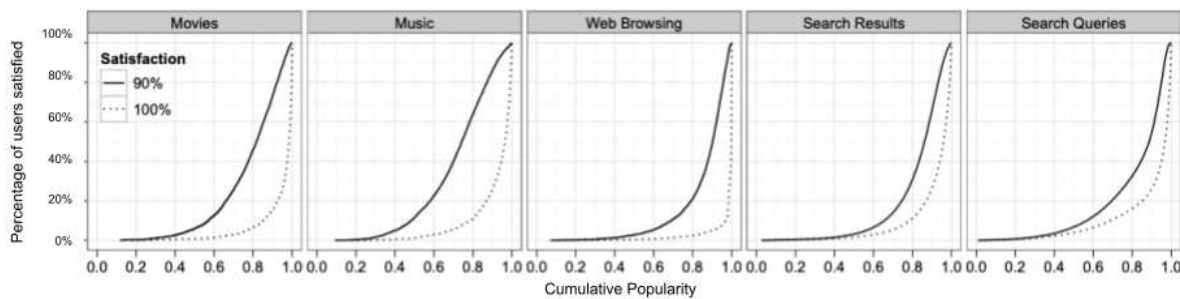


Fig. 2: Percentage of users whose needs are met 90% and 100% of the time, as a function of the cumulative popularity of the top items. (source: Goel et al (2010))

The findings of Goel et al (2020) that long tail availability boosts head sales is also reflected in the automobile market. Referring again to Fig. 1, we see that several manufacturers carry mass market brands in the head of the distribution, as well as niche luxury brands in the long tail, e.g., Toyota & Lexus, Ford & Lincoln, Chevrolet & Cadillac.

In summary, there is good reason to believe that human needs favor long tail distributions, and small increases in coverage along the long tail can result in disproportionately large increases in the satisfaction of the need.

The Long Tail of Labor

If human needs are long-tailed, it follows that services occupations involved in meeting those needs are likely long-tailed as well. At the same time, manufacturing occupations involved in products that meet long-tailed needs may not themselves be long-tailed. For example, although automobiles are a long-tailed need, manufacturing occupations assembling automobiles may not be long-tailed, though repairing and servicing automobiles may be long-tailed. In our context, we are interested in the automation of white-collar occupations such as lawyers, accountants, software engineers etc.

A white-collar occupation typically involves a range of tasks, where each user’s needs is a distribution on the tasks and the occupation serves the aggregate of the user distributions. Some tasks are routine and popular, i.e. the head, while others are niche along the tail. Fig. 3 is an adaptation of Fig. 2, depicting the

tasks involved in such a service. Fig. 3(a) shows the baseline productivity of a human worker performing the service, where he or she is capable of performing 80% of the tasks satisfying ~50% of the users 90% of the time. Fig. 3(b) shows the same worker assisted by an automation technology “Automation-1” that offloads the first 30% of tasks, freeing up the human worker to go deeper into the long tail and perform 90% of the tasks in total, satisfying 80% of the users 90% of the time. Measuring output in terms of the percentage of users satisfied at the 90% level, automation accounts for ~2% of the output, and labor for the remaining ~78%. Hence the worker’s output jumps from ~50% of users without automation in Fig 3(a), to ~78% of users with automation in Fig 3(b), boosting output by $\sim(78-50)/50 = \sim56\%$, potentially enabling him to earn higher wages. Fig. 3(c) shows the same worker assisted by a different automation technology “Automation-2” that offloads both common tasks and some long-tail tasks, freeing up the worker to go even deeper into the long-tail and perform 100% of the tasks in total. In Fig. 3(c), automation’s output is 20% of users satisfied at the 90% level, the worker’s output is $(100-20) = 80\%$ of users satisfied at the 90% level, resulting in a gain of $(80-78)/78 = \sim2.6\%$ compared Fig. 3(b). Further automation reduces the worker’s output, and the worker’s wages may begin to drop. In the limiting case, automation can squeeze out the human worker entirely by covering all of the distribution.

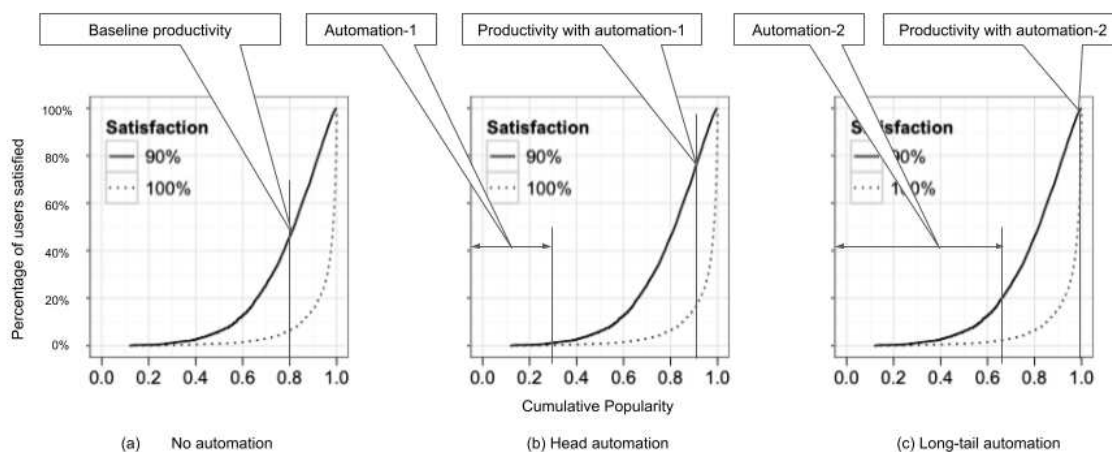


Fig. 3: The impact of automation on long-tailed services

In short, technologies that cover substantial portions of the long tail can squeeze wages or displace humans entirely, while technologies that automate the head may improve human productivity.

Case Study 1: Switchboard Operators

Fixed line telephone calls follow a long tail distribution, where local area calls are most popular, i.e. the head of the distribution, followed by long-distance calls, and finally international calls to increasingly infrequent locations in the long tail. In the early days of telephony, companies used telephone switchboards where human operators connected calls manually by inserting phone plugs into jacks pairing callers and recipients.

When automated local exchanges were introduced, it became possible for callers to “direct dial” other phones in their same local area network. Operators were still required for long distance and international calls along the long tail. Then, area codes were introduced and long-distance calling did not require operator assistance. Finally, international calling with country codes was introduced, eliminating

switchboard operators entirely. In summary, automation covered all of the long tail distribution, displacing human labor altogether.

Whilst automation eliminated switchboard operators, new jobs appeared in the manufacture of the automated exchange equipment. Historically, this has been the case with automation, as noted in Acemoglu and Restrepo (2019), Autor et al. (2022). In our present context, we examine the worst case effect of automation technologies, where a few humans may develop AI technologies that impact a vast number of jobs, with minimal creation of new jobs.

Case Study 2: Accountants

Over the past twenty years, the widespread availability of personal computers made it possible to automate accounting using desktop and cloud software. A large number of products are available for such purpose, ranging from spreadsheets, to tax software, to bookkeeping software, and ever more sophisticated client-server and cloud applications. Even so, between 2000 and 2022 the number of accountants in the US has grown ~35%, as shown in Fig. 4. For comparison, during the same period, the US population and real GDP grew ~18% and ~55% respectively.



Fig. 4: Number of accountants in the US workforce, 2000-2022. (source: [Fred](#))

Fig. 4 suggests accounting software has augmented human accountants by automating some of the common tasks at the head of the workload distribution, but has not displaced accountants altogether. If accounting software was operating in the head as in Fig. 3(b), the productivity and real wages of accountants would have been stable or increased over the same period. Fig. 5 shows that such is indeed the case, and the real wages of accountants increased from 2000 to 2022.

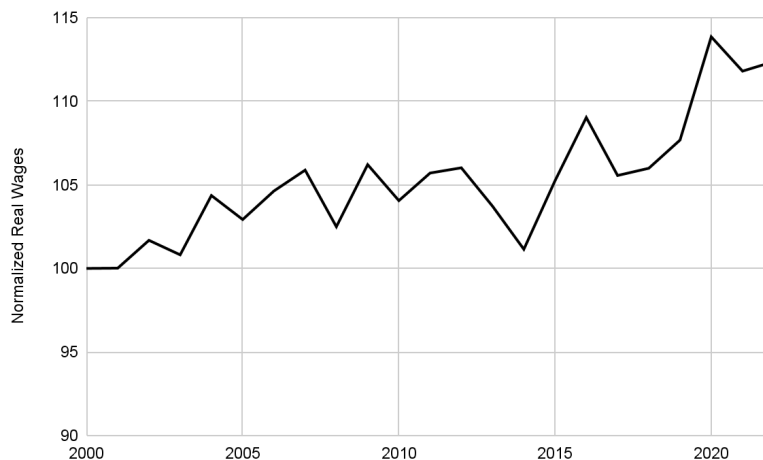


Fig. 5: Normalized real wages of accountants in the US workforce (source: [Fred](#))

The phenomenon of Fig. 4 where automation can increase labor demand initially is visible in other historical economic data, e.g., Bessen (2019). Automation can lower prices, spurring elastic demand for more variants of the output, i.e. the long tail. For example, this is evident in the US automobile market, which transitioned from ~50% market share for the Model T, to the long tail distribution of Fig.1, as the market grew.

In contrast to Figs. 4 and 5, Frey and Osborne (2013) predicted that with 98% probability, accounting jobs will be automated. Subsequently, Arntz, Gregory and Zierahn (2017) validate the phenomenon of Fig. 3(b) on historical economic data:

In light of rapid advances in the fields of Artificial Intelligence (AI) and robotics, many scientists discuss the potentials of new technologies to substitute for human labour. Fuelling the economic debate, various empirical assessments suggest that up to half of all jobs in western industrialized countries are at risk of automation in the next 10 to 20 years. This paper demonstrates that these scenarios are overestimating the share of automatable jobs by neglecting the substantial heterogeneity of tasks within occupations as well as the adaptability of jobs in the digital transformation. To demonstrate this, we use detailed task data and show that, when taking into accounting the spectrum of tasks within occupations, the automation risk of US jobs drops, ceteris paribus, from 38% to 9%.

The Impact of GPT on Labor

We now consider the impact of GPT on labor. Will GPT displace humans and squeeze wages, or will GPT augment humans and increase wages?

Bubeck et al (2023) present a detailed analysis of GPT's strengths and weaknesses as compared to human intelligence. However, for our purposes it suffices to ask the simple question: how does GPT perform on the long tail? We provide empirical evidence that popular GPTs struggle even on relatively simple long tail tasks.

Long Tail Test 1: People

We prompted ChatGPT-3 at chat.openai.com with the names of people in roughly decreasing order of prominence as below. We started with “Barack Obama,” and ChatGPT responded with a brief biography that is factually correct.

Prompt: Barack Obama

ChatGPT: Barack Obama is an American politician and attorney who served as the 44th President of the United States from 2009 to 2017. He was born on August 4, 1961, in Honolulu, Hawaii, and grew up in a multicultural family with a Kenyan father and an American mother. Before becoming president, Obama served as a U.S. Senator from Illinois from 2005 to 2008. During his presidency, he implemented a number of significant policy initiatives, including the Affordable Care Act (also known as Obamacare), the Dodd-Frank Wall Street Reform and Consumer Protection Act, and the American Recovery and Reinvestment Act, which was designed to address the Great Recession. He also signed landmark legislation such as the Lilly Ledbetter Fair Pay Act and the repeal of "Don't Ask, Don't Tell," which barred openly gay individuals from serving in the military. Obama was awarded the Nobel Peace Prize in 2009 for his efforts to strengthen international diplomacy and cooperation between peoples. He is also known for his inspirational speeches, including his "Yes We Can" speech during his 2008 campaign, which became a rallying cry for his supporters. After leaving office, Obama has continued to be active in public life, advocating for issues such as voting rights, climate change, and racial justice.

We then tried “Sam Altman,” who is the CEO of OpenAI, the organization that produces ChatGPT. The errant responses are highlighted.

Prompt: sam altman

ChatGPT: Sam Altman is a technology entrepreneur, investor, and startup advisor based in Silicon Valley, California. He is best known for his role as the president of startup accelerator Y Combinator, which has helped launch companies like Airbnb, Dropbox, and Stripe. Altman also co-founded Loopt, a location-based social networking app, and **was the CEO** of the artificial intelligence research company OpenAI.

Altman is a prominent figure in the tech industry and has been recognized for his contributions to entrepreneurship and innovation. In addition to his work with Y Combinator and OpenAI, he serves on the **board of directors for several companies, including Airbnb, Reddit, and Okta**. He is also an active angel investor and has backed numerous successful startups over the years.

The response says Altman “was the CEO” rather than “is the CEO” of OpenAI. The response also says Altman is on the board of directors of AirBnb and Okta, disagreeing with both companies’ web sites.

Finally, we tried “Prabhakar Raghavan,” who heads Google’s search and GPT products.

Prompt: Prabhakar Raghavan

ChatGPT: Prabhakar Raghavan is an Indian-American computer scientist and the Senior Vice President of Google responsible for Search, Assistant, and Ads. He was born in Chennai, India, in 1963, and received his Bachelor's degree in Mathematics from the Indian Institute of Technology, Kanpur in 1985. He then went on to obtain his Ph.D. in computer science from the University of California, Berkeley in 1989.

Raghavan is a renowned expert in the field of search and information retrieval, and has made significant contributions to the development of algorithms and techniques for web search. He is also known for his work on machine learning and data mining, and has authored several influential books and research papers in these areas.

Prior to joining Google in 2018, Raghavan served as the Vice President of Engineering at Google- owned online advertising company, DoubleClick. He has also held positions at IBM Almaden Research Center, Verity, and Yahoo! Research.

Raghavan has been recognized with numerous awards for his contributions to computer science, including the ACM SIGIR Gerard Salton Award, the ACM SIGMOD Edgar F. Codd Innovations Award, and the IEEE Technical Achievement Award.

The response was factually incorrect in multiple places as compared to Wikipedia and other reliable sources. (1) Raghavan graduated from the Indian Institute of Technology, Madras rather than Kanpur (2) Raghavan graduated from UC Berkeley in 1987 rather than 1989. (3) Raghavan did not work for DoubleClick (4) The list of winners of the ACM SIGIR Gerard Salton Award does not list Raghavan. (5) The list of winners of the ACM SIGMOD Edgar F. Codd Innovations Award does not list Raghavan. (6) We could not find any evidence of Raghavan’s IEEE Technical Achievement Award.

We also tried the same tests on Google’s GPT at bard.google.com, errant responses highlighted below.

Prompt: barack obama

Bard: Barack Obama was the 44th president of the United States. He was born in Honolulu, Hawaii on August 4, 1961. His mother, Ann Dunham, was from Kansas, and his father, Barack Obama Sr., was from Kenya. Obama's parents divorced when he was two years old, and he was raised by his mother and her parents.

Obama attended Columbia University and Harvard Law School. After graduating from law school, he worked as a community organizer in Chicago. He then worked as a civil rights attorney and an academic. In 1996, he was elected to the Illinois State Senate. In 2004, he was elected to the United States Senate.

In 2008, Obama was elected president of the United States. He was the first African American to be elected president. Obama served two terms as president. During his presidency, he oversaw the end of the Iraq War, the killing of Osama bin Laden, and the passage of the Affordable Care Act. Obama left office in 2017. He and his wife, Michelle, live in Washington, D.C. They have two daughters, Malia and Sasha.

Prompt: sam altman

Bard: Sam Altman is an American entrepreneur, investor, and programmer. He is the CEO of OpenAI, Loopt, and Worldcoin. Altman was previously the president of Y Combinator and briefly as the CEO of Reddit.

Altman was born in Chicago, Illinois in 1985. He attended John Burroughs School in St. Louis, Missouri and then Stanford University, where he studied computer science. After graduating from Stanford, Altman worked as a software engineer at Google.

In 2010, Altman co-founded Loopt, a location-based social networking app. Loopt was acquired by Zillow in 2012 for \$10 million. Altman then co-founded OpenAI, a non-profit research company focused on artificial intelligence. OpenAI is funded by Altman, Elon Musk, Peter Thiel, and other Silicon Valley investors.

In 2016, Altman was appointed president of Y Combinator, a startup accelerator. Altman left Y Combinator in 2018 to become the CEO of Reddit. Reddit is a social news aggregation, content rating, discussion website. Altman resigned from Reddit in 2019.

Altman is a controversial figure in the tech industry. He has been criticized for his views on artificial intelligence and his involvement in the startup community. However, Altman is also seen as a visionary entrepreneur who is pushing the boundaries of what is possible in technology.

Prompt: prabhakar raghavan

Bard: I am still working to learn more languages, so I can't do that just yet. Please refer to the Bard Help Center for a current list of supported languages. Is there anything else you'd like my help with?

For humans, the biographies of people is not a difficult task, and yet popular GPTs struggle entering the long tail.

Long Tail Test 2: Math.

Prompt: what is pi to 15 digits

ChatGPT: Pi to 15 digits is 3.141592653589793.

Prompt: what is the 3rd digit after the decimal of pi

ChatGPT: The 3rd digit after the decimal of pi is 1.

Prompt: what is the fifth digit after the decimal of pi

ChatGPT: The fifth digit after the decimal of pi is 9.

Prompt: what is the 7th digit after the decimal of pi

ChatGPT: The 7th digit after the decimal of pi is 5.

Prompt: what is the 9th digit after the decimal of pi

ChatGPT: The 9th digit after the decimal of pi is 3.

Prompt: what is the 12th digit after the decimal of pi

ChatGPT: The 12th digit after the decimal of pi is 5.

Prompt: what is the 14th digit after the decimal of pi

ChatGPT: The 14th digit after the decimal of pi is 6.

In the foregoing, the responses are incorrect on the 7th, 12th and 14 digits as highlighted.

We repeated the tests on Google's Bard GPT as below.

Prompt: what is pi to 15 digits

Bard: Pi to 15 digits is 3.14159265358979. It is an irrational number, which means that its decimal representation never ends or repeats. Pi is the ratio of the circumference of a circle to its diameter. It is one of the most important mathematical constants, and it has applications in many fields, including geometry, physics, and engineering.

Prompt: what is the 3rd digit after the decimal of pi

Bard: The third digit after the decimal point of pi is 1.

Prompt: what is the fifth digit after the decimal of pi

Bard: The fifth digit after the decimal point of pi is 9.

Prompt: what is the 7th digit after the decimal of pi

Bard: There is no 7th digit after the decimal point of pi. The decimal representation of pi is infinite, so there is no last digit.

Prompt: what is the 9th digit after the decimal of pi

Bard: There is no 9th digit after the decimal of pi. As mentioned above, the decimal representation of pi is infinite, so there is no last digit.

Prompt: what is the 12th digit after the decimal of pi

Bard: There is no 12th digit after the decimal of pi. As mentioned above, the decimal representation of pi is infinite, so there is no last digit.

The first 12 digits of pi are 3.14159265358979323846264338327950288419716939937510

Prompt: what is the 14th digit after the decimal of pi

Bard: As pi is an irrational number, its decimal representation never ends or repeats. Therefore, there is no 14th digit after the decimal of pi.

For humans, these tasks are not difficult. Yet, popular GPTs struggle beyond the first few, and are erratic entering the long tail.

Long Tail Test: Search+GPT

Some GPTs are connected to search engines, combining the human insight of search rank with the capabilities of a GPT, e.g. Microsoft Bing with GPT-4. Such systems can summarize the top few results from a search engine when appropriate. For the People test, Bing summarized the top few lines of Wikipedia correctly for each individual. And for the Math test, Bing found web sites that specifically provided the "nth digit after the decimal of pi." However, Bing struggled on tasks for which the answer was not readily available in search results. For example, we ran the Math test replacing "pi" with "pi/3."

Prompt: pi/3

Bing: The value of pi/3 is approximately 1.0471975511965976 .

Prompt: 1st digit after the decimal of pi/3

Bing: The first digit after the decimal point of pi/3 is 1

Prompt: 4th digit after the decimal of pi/3

Bing: The fourth digit after the decimal point of pi/3 is 3

Prompt: 7th digit after the decimal of pi/3

Bing: The seventh digit after the decimal point of pi/3 is 9

Prompt: 9th digit after the decimal of pi/3

Bing: The ninth digit after the decimal point of pi/3 is 7

Long Tail Test: Medicine

Harskamp and De Clercq (2023) tested ChatGPT on medical questions of varying complexity:

We found that ChatGPT correctly answered 74% of the trivia questions, with slight variation in accuracy in the domains coronary artery disease (80%), pulmonary and venous thrombotic embolism (80%), atrial fibrillation (70%), heart failure (80%) and cardiovascular risk management (60%). In the case vignettes, ChatGPT's response matched in 90% of the cases with the actual advice given. In more complex cases, where physicians (general practitioners) asked other physicians (cardiologists) for assistance or decision support, ChatGPT was correct in 50% of cases, and often provided incomplete or inappropriate recommendations when compared with expert consultation.

In short, popular GPTs are not yet capable of straightforward long-tail tasks. In their present form, we believe these technologies may automate head tasks, freeing up humans to focus on more complex long-tail tasks and improve productivity, but are unlikely to squeeze wages or displace humans entirely. Our observations are experimentally validated by Noy and Zhang (2023), in the context of mid-level writing tasks:

ChatGPT mostly substitutes for worker effort rather than complementing worker skills, and restructures tasks towards idea-generation and editing and away from rough-drafting. Exposure to ChatGPT increases job satisfaction and self-efficacy and heightens both concern and excitement about automation technologies.

We note that GPTs are improving rapidly, and future generations of the technology might well perform on the long tail. We also note that the long tail can be stubbornly elusive, as evidenced by self-driving cars which remain erratic on the long tail despite early promise and vast investment. Chafkin (2022) writes:

Over the course of more than a decade, flashy demos from companies including Google, GM, Ford, Tesla, and Zoox have promised cars capable of piloting themselves through chaotic urban landscapes, on highways, and in extreme weather without any human input or oversight. The companies have suggested they're on the verge of eliminating road fatalities, rush-hour traffic, and parking lots, and of upending the \$2 trillion global automotive industry...there's an entire social media genre featuring self-driving cars that become hopelessly confused. When the results are less serious, they can be funny as hell. In one example, a Waymo car gets so flummoxed by a traffic cone that it drives away from the technician sent out to rescue it. In another, an entire fleet of modified Chevrolet Bolts show up at an intersection and simply stop, blocking traffic with a whiff of Maximum Overdrive. In a third, a Tesla drives, at very slow speed, straight into the tail of a private jet....While the industry's biggest names continue to project optimism, the emerging consensus is that the world of robo-taxis isn't just around the next unprotected left—that we might have to wait decades longer, or an eternity.

Summary

We set out to identify a test to predict the impact on labor of emerging automation technologies such as GPT. Towards that goal, we observed that human needs favor long tail distributions, demanding the same from the products and services that fulfill those needs. Technologies that address a portion of the distribution, typically the head, free up human labor to focus on more complex tasks in the long tail thereby improving productivity and potentially lifting wages. For example, despite the large-scale adoption of personal computers and accounting software, the number of accountants and their real wages increased in recent decades. In contrast, technologies that cover substantial portions of the long tail can squeeze wages or displace humans entirely. For example, automated telephone exchanges completely replaced human telephone operators, since they addressed the problem across its full distribution. With this in mind, we proposed a long tail test for automation technologies to predict their impact on labor. We found that popular GPTs perform poorly on such tests in that they are erratic on straightforward long tail tasks, hence absent breakthroughs, will augment human productivity rather than cause mass displacement of human labor. Going forward, we believe that to have a broad impact on displacing or devaluing human labor, AI must at least be capable of long-tail tasks that humans perform with ease.

Acknowledgements

Thanks to F. Baskett, J. Jawahar, G. Johnson, L. Katz, R. Krishnan, A.M. Salomons, K. Suresh and P. Tadepalli.

References

- Fowler, Gregory, A. (2023): “Say what, Bard? What Google’s new AI gets right, wrong and weird”
<https://www.washingtonpost.com/technology/2023/03/21/google-bard/>
- Roose, Kevin, (2023): “A Conversation With Bing’s Chatbot Left Me Deeply Unsettled”
<https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html>
- Korinek, Anton and Stiglitz, Joseph, (2020): “Steering Technological Progress,”
https://conference.nber.org/conf_papers/f143989.pdf
- Brynjolfsson, Erik, (2022): “The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence,”
<https://digitaleconomy.stanford.edu/news/the-turing-trap-the-promise-peril-of-human-like-artificial-intelligence/>
- Acemoglu, Daron, (2022): “The Harms of AI,” <https://economics.mit.edu/files/21848>
- Autor, D., Mindell, D., and Reynolds, E. (2020). “The Work of the Future: Building Better Jobs in an Age of Intelligent Machines” Cambridge: MIT Press.
<https://workofthefuture.mit.edu/wp-content/uploads/2021/01/2020-Final-Report4.pdf>

Basu, Kaushik, (2022), “An inclusive future? Technology, New Dynamics and Policy Challenges”
Brookings Institute,
https://www.brookings.edu/wp-content/uploads/2022/05/Inclusive-future_Technology-new-dynamics-policy-challenges.pdf

Acemoglu, D., and Restrepo, P., (2019). “Automation and New Tasks: How Technology Displaces and Reinstates Labor.” *Journal of Economic Perspectives*—Volume 33, Number 2.

Autor, D., Chin, C., Salomons, A.M., & Seegmiller, B., (2022), “New Frontiers: The Origins and Content of New Work, 1940–2018,” <https://www.nber.org/papers/w30389>

Eloundou, T., Manning, M., Mishkin, P., and Rock, D. (2023). “GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models,” <https://arxiv.org/pdf/2303.10130.pdf>

Anderson, C., (2004), “The long tail.” *Wired Magazine*, 12(10):170–177, 2004.

Anderson, C., (2006), “The Long Tail: Why the Future of Business is Selling Less of More.” Hyperion.

Brynjolfsson, E., Hu, Y.J., and Smith, M.D., (2007), “Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers.” MIT Center for Digital Business Working Paper.

Goel, S., Broder, A, Gabrilovich, E., Pang, B., (2010). “Anatomy of the long tail: ordinary people with extraordinary tastes.” *Proceedings of the third ACM international conference on Web search and data mining*.

Bessen, J.E., (2019), “Automation and Jobs: When Technology Boosts Employment”
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2935003

Frey, C. and M.A. Osborne (2013). “The Future of Employment: How Susceptible are Jobs to Computerization?”
https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

Arntz, M., Gregory, T. & Zierahn, U., “Revisiting the Risk of Automation,” *July 2017 Economics Letters* 159:157-160

Bubeck, S., Chandrasekharan, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y-T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribiero, M.T., Zhang, Y., (2023). “Sparks of Artificial General Intelligence: Early experiments with GPT-4.” <https://arxiv.org/pdf/2303.12712.pdf>

Harskamp, R.E., and De Clercq, L., (2023), “Performance of ChatGPT as an AI-assisted decision support tool in medicine: a proof-of-concept study for interpreting symptoms and management of common cardiac conditions (AMSTELHEART-2),”
<https://www.medrxiv.org/content/10.1101/2023.03.25.23285475v1.full.pdf>

Noy, S. and Zhang, W., (2023). “Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence,” https://economics.mit.edu/sites/default/files/inline-files/Noy_Zhang_1.pdf

Chafkin, (2022), “Even After \$100 Billion, Self-Driving Cars Are Going Nowhere”
<https://www.bloomberg.com/news/features/2022-10-06/even-after-100-billion-self-driving-cars-are-going-nowhere>