GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models

Tyna Eloundou[1], Sam Manning[1,2], Pamela Mishkin*[1], and Daniel Rock[3]

[1]OpenAI
[2]OpenResearch
[3]University of Pennsylvania

March 27, 2023

Abstract

We investigate the potential implications of large language models (LLMs), such as Generative Pre-trained Transformers (GPTs), on the U.S. labor market, focusing on the increased capabilities arising from LLM-powered software compared to LLMs on their own. Using a new rubric, we assess occupations based on their alignment with LLM capabilities, integrating both human expertise and GPT-4 classifications. 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. This finding implies that LLM-powered software will have a substantial effect on scaling the economic impacts of the underlying models. We conclude that LLMs such as GPTs exhibit traits of general-purpose technologies, indicating that they could have considerable economic, social, and policy implications.

1 Introduction

As shown in Figure 1, recent years, months, and weeks have seen remarkable progress in the field of generative AI and large language models (LLMs). While the public often associates LLMs with various iterations of the Generative Pre-trained Transformer (GPT), LLMs can be trained using a range of architectures, and are not limited to transformer-based models (Devlin et al., 2019). LLMs can process and produce various forms of sequential data, including assembly language, protein sequences and chess games, extending beyond natural language applications alone. In this paper, we use LLMs and GPTs somewhat interchangeably, and specify in our rubric that these should be considered similar to the GPT-family of models available via ChatGPT or the OpenAI Playground (which at the time of labeling included models in the GPT-3.5 family but not in the GPT-4 family). We examine LLMs with text- and code-generating abilities, use the term "generative AI" to additionally include modalities such as images or audio, and use "LLM-powered software" to cover tools built on top of LLMs or that combine LLMs with other generative AI models.

*Corresponding author (pamela@openai.com). Authors contributed equally and are listed alphabetically.

Group	Occupations with highest exposure	% Exposure
Human α	Interpreters and Translators	76.5
	Survey Researchers	75.0
	Poets, Lyricists and Creative Writers	68.8
	Animal Scientists	66.7
	Public Relations Specialists	66.7
Human β	Survey Researchers	84.4
	Writers and Authors	82.5
	Interpreters and Translators	82.4
	Public Relations Specialists	80.6
	Animal Scientists	77.8
Human ζ	Mathematicians	100.0
	Tax Preparers	100.0
	Financial Quantitative Analysts	100.0
	Writers and Authors	100.0
	Web and Digital Interface Designers	100.0
	Humans labeled 15 occupations as "fully exposed."	
Model α	Mathematicians	100.0
	Correspondence Clerks	95.2
	Blockchain Engineers	94.1
	Court Reporters and Simultaneous Captioners	92.9
	Proofreaders and Copy Markers	90.9
Model β	Mathematicians	100.0
	Blockchain Engineers	97.1
	Court Reporters and Simultaneous Captioners	96.4
	Proofreaders and Copy Markers	95.5
	Correspondence Clerks	95.2
Model ζ	Accountants and Auditors	100.0
	News Analysts, Reporters, and Journalists	100.0
	Legal Secretaries and Administrative Assistants	100.0
	Clinical Data Managers	100.0
	Climate Change Policy Analysts	100.0
	The model labeled 86 occupations as "fully exposed."	
Highest variance	Search Marketing Strategists	14.5
	Graphic Designers	13.4
	Investment Fund Managers	13.0
	Financial Managers	13.0
	Insurance Appraisers, Auto Damage	12.6

Table 4: Occupations with the highest exposure according to each measurement. The final row lists the occupations with the highest σ^2 value, indicating that they had the most variability in exposure scores. Exposure percentages indicate the share of an occupation's task that are exposed to GPTs ($\boldsymbol{\alpha}$) or GPT-powered software ($\boldsymbol{\beta}$ and $\boldsymbol{\zeta}$), where exposure is defined as driving a reduction in time it takes to complete the task by at least 50% (see exposure rubric A.1). As such, occupations listed in this table are those where we estimate that GPTs and GPT-powered software are able to save workers a significant amount of time completing a large share of their tasks, but it does not necessarily suggest that their tasks can be fully automated by these technologies.

smooth the transition to an economy with increasingly widespread LLM adoption, prior work such as (Autor et al., 2022b) has articulated several important directions for US policy related to education, worker training, reforms to safety net programs, and more.

6.3 Limitations and Future Work

In addition to those discussed above, we highlight some particular limitations of this work that warrant further investigation. Primarily, our focus on the United States restricts the generalizability of our findings to other nations where the adoption and impact of generative models may differ due to factors such as industrial organization, technological infrastructure, regulatory frameworks, linguistic diversity, and cultural contexts. We hope to address this limitation by extending the study's scope and by sharing our methods so other researchers can build on them.

Subsequent research efforts should consider two additional studies: one exploring LLM adoption patterns across various sectors and occupations, and another scrutinizing the actual capabilities and limitations of state-of-the-art models in relation to worker activities beyond the scope of our exposure scores. For example, despite recent advances in multimodal capabilities with GPT-4, we did not consider vision capabilities in the α ratings on direct LLMs-exposure (OpenAI, 2023b). Future work should consider the impact of such capability advances as they unfold. Furthermore, we acknowledge that there may be discrepancies between theoretical and practical performance, particularly in complex, open-ended, and domain-specific tasks.

7 Conclusion

In conclusion, this study offers an examination of the potential impact of LLMs on various occupations and industries within the U.S. economy. By applying a new rubric for understanding LLM capabilities and their potential effects on jobs, we have observed that most occupations exhibit some degree of exposure to LLMs, with higher-wage occupations generally presenting more tasks with high exposure. Our analysis indicates that approximately 19% of jobs have at least 50% of their tasks exposed to LLMs when considering both current model capabilities and anticipated LLM-powered software.

Our research aims to highlight the general-purpose potential of LLMs and their possible implications for US workers. Previous literature demonstrates the impressive improvements of LLMs to date (see 2.1). Our findings confirm the hypothesis that these technologies can have pervasive impacts across a wide swath of occupations in the US, and that additional advancements supported by LLMs, mainly through software and digital tools, can have significant effects on a range of economic activities. However, while the technical capacity for LLMs to make human labor more efficient appears evident, it is important to recognize that social, economic, regulatory, and other factors will influence actual labor productivity outcomes. As capabilities continue to evolve, the impact of LLMs on the economy will likely persist and increase, posing challenges for policymakers in predicting and regulating their trajectory.

Further research is necessary to explore the broader implications of LLM advancements, including their potential to augment or displace human labor, their impact on job quality, impacts on inequality, skill development, and numerous other outcomes. By seeking to understand the capabilities and potential effects of LLMs on the workforce, policymakers and stakeholders can make more informed decisions to navigate the complex landscape of AI and its role in shaping the future of work.

7.1 LLM Conclusion (GPT-4's Version)

Generative Pre-trained Transformers (GPTs) generate profound transformations, garnering potential technological growth, permeating tasks, greatly impacting professions. This study probes GPTs' potential trajectories, presenting a groundbreaking rubric to gauge tasks' GPT exposure, particularly in the U.S. labor market.

E Occupations Without Any Exposed Tasks

Occupations with no labeled exposed tasks
Agricultural Equipment Operators
Athletes and Sports Competitors
Automotive Glass Installers and Repairers
Bus and Truck Mechanics and Diesel Engine Specialists
Cement Masons and Concrete Finishers
Cooks, Short Order
Cutters and Trimmers, Hand
Derrick Operators, Oil and Gas
Dining Room and Cafeteria Attendants and Bartender Helpers
Dishwashers
Dredge Operators
Electrical Power-Line Installers and Repairers
Excavating and Loading Machine and Dragline Operators, Surface Mining
Floor Layers, Except Carpet, Wood, and Hard Tiles
Foundry Mold and Coremakers
Helpers–Brickmasons, Blockmasons, Stonemasons, and Tile and Marble Setters
Helpers–Carpenters
Helpers–Painters, Paperhangers, Plasterers, and Stucco Masons
Helpers–Pipelayers, Plumbers, Pipefitters, and Steamfitters
Helpers–Roofers
Meat, Poultry, and Fish Cutters and Trimmers
Motorcycle Mechanics
Paving, Surfacing, and Tamping Equipment Operators
Pile Driver Operators
Pourers and Casters, Metal
Rail-Track Laying and Maintenance Equipment Operators
Refractory Materials Repairers, Except Brickmasons
Roof Bolters, Mining
Roustabouts, Oil and Gas
Slaughterers and Meat Packers
Stonemasons
Tapers
Tire Repairers and Changers
Wellhead Pumpers

Table 11: All 34 occupations for which none of our measures labeled any tasks as exposed.

References

Abid, A., Farooqi, M., and Zou, J. (2021). Persistent anti-muslim bias in large language models. In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, AIES ’21, page 298–306, New York, NY, USA. Association for Computing Machinery.