

How will Language Modelers like ChatGPT Affect Occupations and Industries?

Ed Felten (Princeton)

Manav Raj (University of Pennsylvania)

Robert Seamans (New York University)

1 March 2023

Abstract: Recent dramatic increases in AI language modeling capabilities has led to many questions about the effect of these technologies on the economy. In this paper we present a methodology to systematically assess the extent to which occupations, industries and geographies are exposed to advances in AI language modeling capabilities. We find that the top occupations exposed to language modeling include telemarketers and a variety of post-secondary teachers such as English language and literature, foreign language and literature, and history teachers. We find the top industries exposed to advances in language modeling are legal services and securities, commodities, and investments.

Keywords: artificial intelligence, ChatGPT, language modeling, occupation, technology

1. Introduction

Artificial Intelligence (AI) will likely affect the economy in many ways, potentially boosting economic growth and changing the way people work and play. The effect of AI on work will likely be multi-faceted. In some cases, AI may substitute for work previously done by humans, and in other cases AI may complement work done by humans. The effect on work will likely also vary across industries. Recent research by Goldfarb et al (2020) document that adoption of AI is relatively high in some industries such as information technology and finance, but low in others such as health care and construction. Moreover, trying to understand how AI will affect work is like trying to hit a moving target because the capabilities of AI are still advancing.

A prominent example of how AI capabilities continue to advance are the recent improvements in AI language modeling. In particular, ChatGPT, a language modeler released by Open AI in late 2022, has garnered a huge amount of attention and controversy. Some worry about the negative effects of tools like ChatGPT on jobs, as in the *New York Post* article headlined “ChatGPT could make these jobs obsolete: ‘The wolf is at the door.’”¹ Others see practical and commercial promise from language modeling. For example, Microsoft announced a \$10 billion partnership with Open AI and has linked ChatGPT with its Bing search engine.² Google felt compelled to demonstrate its own language modeler, Bard, but mistakes during the demonstration led Google’s stock price to drop 7%.³ ChatGPT has been banned by J.P. Morgan.⁴ However, at present, most of this is speculation.

In order to better understand how language modelers such as ChatGPT will affect occupations, industries and geographies, we use a methodology developed by Felten et al (2018, 2021). Felten et al created the AI Occupational Exposure (AIOE) measure and used this measure to identify which occupations, industries and geographies are most exposed to AI. In this paper, we describe how the AIOE approach can be adapted to account for the recent advancement of language modeling.

¹ <https://nypost.com/2023/01/25/chat-gpt-could-make-these-jobs-obsolete/>

² <https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai>

³ <https://www.cnbc.com/2023/02/08/alphabet-shares-slip-following-googles-ai-event-.html>

⁴ <https://www.cbsnews.com/news/chatgpt-jpmorgan-chase-bars-workers-from-using-ai-tool/>

We find that the top occupations affected include telemarketers and a variety of post-secondary teachers such as English language and literature, foreign language and literature, and history teachers. We also find the top industries exposed to advances in language modeling are legal services and securities, commodities, and investments.

This article contributes to several literatures. First, by providing a systematic examination of the effect of language modeling across occupations, industries and geographies, it contributes to a nascent literature on the effects of ChatGPT and other language modelers on the economy (e.g. Agarwal et al., 2022; Zarifhonarvar, 2023). More generally, the article builds on a broader set of literature studying the effect of AI on the economy (Furman and Seamans, 2019; Goldfarb et al., 2019). Second, the article builds on and extends a set of papers that provide systematic methodologies for studying how AI affects occupations (e.g., Brynjolfsson et al, 2018; Frey & Osborne, 2017; Tolan et al., 2021; Webb, 2020). The article specifically builds off and extends the methodology described in Felten et al. (2018, 2021). In so doing, the article demonstrates the flexibility of the original Felten et al methodology; it can be adjusted dynamically to assess the impact of changes in AI capabilities. Finally, the article adds to a large literature on the effect of automating technologies on labor (e.g., Acemoglu et al., 2022; Autor, 2015; Frank et al., 2019; Genz et al., 2021).

The article proceeds as follows. Section 2 describes the AI Occupational Exposure (AIOE) measure developed by Felten et al (2018, 2021). Section 3 extends the AIOE to account for recent advances in language modeling. Section 4 provides results, including listing the top 20 most affected occupations and industries. Section 5 concludes.

2. AI Occupational Exposure Methodology

According to Felten et al (2021), the AI Occupational Exposure (AIOE) is a measure of each occupation's "exposure" to AI. The term "exposure" is used so as to be agnostic as to the effects of AI on the occupation, which could involve substitution or augmentation depending on various factors associated with the occupation itself.

The AIOE measure was constructed by linking 10 AI applications (abstract strategy games, real-time video games, image recognition, visual question answering, image generation, reading comprehension, language modeling, translation, speech recognition, and instrumental track

recognition) to 52 human abilities (e.g., oral comprehension, oral expression, inductive reasoning, arm-hand steadiness, etc) using a crowd-sourced matrix that indicates the level of relatedness between each AI application and human ability. Data on the AI applications come from the Electronic Frontier Foundation (EFF) which collects and maintains statistics about the progress of AI across multiple applications. Data on human abilities comes from the Occupational Information Network (O*NET) database developed by the United States Department of Labor. O*NET uses these 52 human abilities to describe the occupational makeup of each of 800+ occupations that it tracks. Each of 800+ occupations can be thought of as a weighted combination of the 52 human abilities. O*NET uses two sets of weights: prevalence and importance.

Once the 10 AI categories and 52 human abilities are linked through the matrix, the AIOE can then be calculated for each occupation. To do this, first we calculate an ability-level exposure as follows:

$$A_{ij} = \sum_{i=1}^{10} x_{ij} \quad (1)$$

Where i indexes the AI application and j indexes the occupational ability. The ability-level exposure, A , is calculated as the sum of the 10 application-ability relatedness scores, x , as constructed using the matrix of crowd-sourced survey data.

We then calculate the AIOE for each occupation k as follows:

$$AIOE_k = \frac{\sum_{j=1}^{52} A_{ij} \times L_{jk} \times I_{jk}}{\sum_{j=1}^{52} L_{jk} \times I_{jk}} \quad (2)$$

In this equation, i indexes the AI application, j indexes the occupational ability, and k indexes the occupation. A_{ij} represents the ability-level exposure score. We weight the ability-level AI exposure by the ability's prevalence (L_{jk}) and importance (I_{jk}) within each occupation as measured by O*NET by multiplying the ability-level AI exposure by the prevalence and importance scores for that ability within each occupation, scaled so that they are equally weighted.

Felten et al (2021) explain the construction of the AIOE scores in more detail, describe how they can be weighted at the industry level to construct an AI Industry Exposure score, or weighted at the geographic level to construct an AI Geographic Exposure score. They also provide results

from a number of validation exercises and describe a number of ways in which the scores can be used by scholars and practitioners.⁵

3. Language Modeling AI Occupational Exposure

The original AIOE described in Felten et al (2021) explicitly weighted each of the AI applications the same. In order to update the AI Occupational Exposure score to account for advances in Language Modeling we modify equation (1) as follows.

$$A_{ij} = \sum_{i=1}^{10} \alpha_i x_{ij} \quad (3)$$

Where i indexes the AI application and j indexes the occupational ability. The ability-level exposure, A , is calculated as the weighted sum of the 10 application-ability relatedness scores, x , as constructed using the matrix of crowd-sourced survey data. α_i is the weight placed on each application i . The weights used in Felten et al (2021) set α_i equal to 1 for each application i .

Next, we set α_i equal to 0 for every AI application except for language modeling, which retains a weight of 1. This then constructs an ability-level exposure measure that only “counts” the value of abilities that are related to language modeling. We then proceed to calculate the $AIOE_k$ for each occupation k using this new “language modeling” weighted A_{ij} . The resulting $AIOE_k$ therefore captures the extent to which each occupation is exposed to advances in language modeling due to AI. A complete list of the occupations and their resulting AIOE language modeling score are listed in an appendix.

The resulting scores are highly correlated with the original AIOE scores (correlation coefficient: 0.979). This can be seen in Figure 1 which plots the original AIOE score and the new language modeling adjusted AIOE score for each occupation.

<< Insert Figure 1 Here >>

4. Results

⁵ The Felten et al (2021) paper is open access and available here: <https://onlinelibrary.wiley.com/doi/full/10.1002/smj.3286> The data and code used to create the AIOE scores described in Felten et al (2021) is available on GitHub: <https://github.com/AIOE-Data/AIOE>

In this section we present and briefly discuss tables of “top 20” occupations and industries exposed to language modeling.

4.1. Top 20 Occupations Exposed to Language Modeling

Table 1 provides the list of top 20 occupations exposed to AI based on the original Felten et al (2021) AI Occupational Exposure (AIOE) measure as well as the top 20 occupations exposed to AI enabled advances in language modeling capabilities.

<< Insert Table 1 Here >>

Some occupations occur in both lists, including “clinical, counseling, and school psychologists”, and “history teachers, postsecondary”. Notably, the language modeling list includes more education-related occupations, indicating that occupations in the field of education are likely to be relatively more impacted by advances in language modeling than other occupations. This accords well with the recent spate of articles around how ChatGPT and other language modeling tools affect the way teachers assign work and detect cheating or could use language modeling tools to develop teaching materials.

Also of interest, the top occupation in the language modeling list is “telemarketer.” One might imagine that human telemarketers could benefit from language modeling being used to augment their work. For example, customer responses can be fed into a language modeling engine in real time and relevant, customer-specific prompts quickly fed to the telemarketer. Or, one might imagine that human telemarketers are substituted with language modeling enabled bots. The potential for language modeling to augment or substitute for human telemarketers work highlights one aspect of the AIOE measure: it measures “exposure” to AI, but whether that exposure leads to augmentation or substitution will depend on specifics of any given occupation.

4.2. Top 20 Industries Exposed to Language Modeling

Table 2 provides the list of 20 industries most exposed to AI based on the original Felten et al. (2021) AI Industry Exposure (AIIE) measure as well as the top 20 industries exposed to AI enabled advances in language modeling capabilities.

<< Insert Table 2 Here >>

As before, we see some similarities in the industries categorized as most exposed to AI based on the original AIOE as well as the version that focuses on advances in language modeling capabilities. For example, “Securities, Commodity Contracts, and Other Financial Investments and Related Activities” is categorized as the most exposed industry using the original AIOE and is the second most exposed industry using the language modeling-focused version of the AIOE. Legal services, insurance and employee benefit funds, and agencies, brokerages, and other insurance related activities are among the top five most exposed industries across both lists.

However, some differences emerge. One salient difference is that the language modeling-focused AIOE suggests a higher exposure to advances in AI within higher education and higher education-adjacent industries. Junior colleges, grantmaking and giving services, and business schools and computer and management training all appear within the top twenty exposed industries.

5. Conclusion

In this paper we present a methodology to systematically assess the extent to which occupations and industries are exposed to advances in AI language modeling capabilities. This methodology relies on the approach described in Felten et al (2021) but adapts it to account for recent advances in language modeling. We find that the top occupations exposed to language modeling include telemarketers and a variety of post-secondary teachers such as English language and literature, foreign language and literature, and history teachers. We also find the top industries exposed to advances in language modeling are legal services and securities, commodities, and investments.

At a broad level, this paper adds to a growing literature studying the effects of AI on labor and work. More specifically, the paper provides a systematic approach for understanding how ChatGPT and other language modelers will affect occupations, industries and geographies. We believe these results will be useful for other scholars as well as practitioners and policymakers.

References

- Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1), S293-S340.
- Agrawal, A., Gans, J., Goldfarb, A. 2022. "ChatGPT and How AI Disrupts Industries" *Harvard Business Review*. Available: <https://hbr.org/2022/12/chatgpt-and-how-ai-disrupts-industries>
- Autor, D. (2015). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives*, 29(3), 3–30.
- Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, 108, 43–47.
<https://doi.org/10.1257/pandp.20181019>
- Felten, E. W., Raj, M., & Seamans, R. (2018). A method to link advances in artificial intelligence to occupational abilities. In *AEA Papers and Proceedings* (Vol. 108, pp. 54-57).
- Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12), 2195-2217.
- Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., ... & Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences*, 116(14), 6531-6539.
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254–280.
<https://doi.org/10.1016/j.techfore.2016.08.019>
- Furman, J., & Seamans, R. (2019). AI and the Economy. *Innovation policy and the economy*, 19(1), 161-191.

Genz, S., Gregory, T., Janser, M., Lehmer, F., & Matthes, B. (2021). How do workers adjust when firms adopt new technologies?. *ZEW-Centre for European Economic Research Discussion Paper*, (21-073).

Goldfarb, A., Gans, J., & Agrawal, A. (2019). *The economics of artificial intelligence: An agenda*. Chicago, IL: University of Chicago Press.

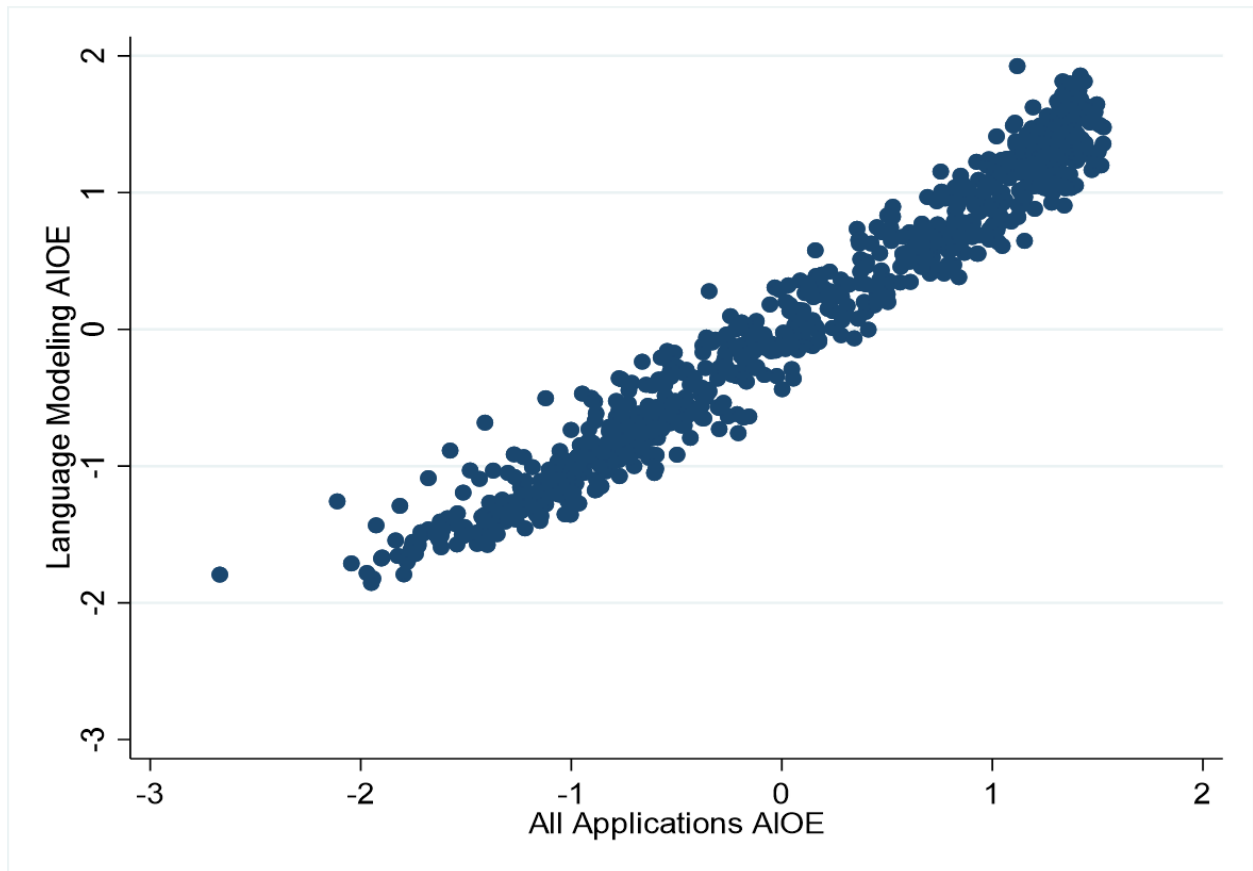
Goldfarb, A., Taska, B. Teodoridis, F. (2020). "Artificial Intelligence in Healthcare? Evidence from Online Job Postings", *AEA Papers and Proceedings*, 110 (5): 400-404

Tolan, S., Pesole, A., Martínez-Plumed, F., Fernández-Macías, E., Hernández-Orallo, J., & Gómez, E. (2021). Measuring the occupational impact of AI: tasks, cognitive abilities and AI benchmarks. *Journal of Artificial Intelligence Research*, 71, 191-236.

Webb, M. (2020). The impact of artificial intelligence on the labor market. *Stanford University working paper*.

Zarifhonorvar, A. 2023. "Economics of ChatGPT: A Labor Market View on the Occupational Impact of Artificial Intelligence" *Indiana University working paper*. Available: <http://dx.doi.org/10.2139/ssrn.4350925>

Figure 1: Comparison between Original AIOE and Language Modeling Adjusted AIOE



Notes: This figure plots the original AIOE score (x-axis) and the new language modeling adjusted AIOE score (y-axis) for each occupation.

Table 1: Top 20 Occupations Exposed to AI, Original and with Language Modeling Adjustment

Rank	Top 20 Occupations from Original AIOE	Top 20 Occupations after Language Modeling Adjustment
1	Genetic Counselors	Telemarketers
2	Financial Examiners	English Language and Literature Teachers, Postsecondary
3	Actuaries	Foreign Language and Literature Teachers, Postsecondary
4	Purchasing Agents, Except Wholesale, Retail, and Farm Products	History Teachers, Postsecondary
5	Budget Analysts	Law Teachers, Postsecondary
6	Judges, Magistrate Judges, and Magistrates	Philosophy and Religion Teachers, Postsecondary
7	Procurement Clerks	Sociology Teachers, Postsecondary
8	Accountants and Auditors	Political Science Teachers, Postsecondary
9	Mathematicians	Criminal Justice and Law Enforcement Teachers, Postsecondary
10	Judicial Law Clerks	Sociologists
11	Education Administrators, Postsecondary	Social Work Teachers, Postsecondary
12	Clinical, Counseling, and School Psychologists	Psychology Teachers, Postsecondary
13	Financial Managers	Communications Teachers, Postsecondary
14	Compensation, Benefits, and Job Analysis Specialists	Political Scientists
15	Credit Authorizers, Checkers, and Clerks	Area, Ethnic, and Cultural Studies Teachers, Postsecondary
16	History Teachers, Postsecondary	Arbitrators, Mediators, and Conciliators
17	Geographers	Judges, Magistrate Judges, and Magistrates
18	Epidemiologists	Geography Teachers, Postsecondary
19	Management Analysts	Library Science Teachers, Postsecondary
20	Arbitrators, Mediators, and Conciliators	Clinical, Counseling, and School Psychologists

Notes: This table lists the top 20 occupations most exposed to AI from the original AIOE (Felten et al., 2021) and the top 20 occupations most exposed to language modeling.

Table 2: Top 20 Industries Exposed to AI, Original and with Language Modeling Adjustment

Rank	Top 20 Industries from Original AIOE	Top 20 Industries after Language Modeling Adjustment
1	Securities, Commodity Contracts, and Other Financial Investments and Related Activities	Legal Services
2	Accounting, Tax Preparation, Bookkeeping, and Payroll Services	Securities, Commodity Contracts, and Other Financial Investments and Related Activities
3	Insurance and Employee Benefit Funds	Agencies, Brokerages, and Other Insurance Related Activities
4	Legal Services	Insurance and Employee Benefit Funds
5	Agencies, Brokerages, and Other Insurance Related Activities	Nondepository Credit Intermediation
6	Nondepository Credit Intermediation	Agents and Managers for Artists, Athletes, Entertainers, and Other Public Figures
7	Other Investment Pools and Funds	Insurance Carriers
8	Insurance Carriers	Other Investment Pools and Funds
9	Software Publishers	Accounting, Tax Preparation, Bookkeeping, and Payroll Services
10	Lessors of Nonfinancial Intangible Assets (except Copyrighted Works)	Business Support Services
11	Agents and Managers for Artists, Athletes, Entertainers, and Other Public Figures	Software Publishers
12	Credit Intermediation and Related Activities (5221 And 5223 only)	Lessors of Nonfinancial Intangible Assets (except Copyrighted Works)
13	Computer Systems Design and Related Services	Business Schools and Computer and Management Training
14	Management, Scientific, and Technical Consulting Services	Credit Intermediation and Related Activities (5221 And 5223 only)
15	Monetary Authorities-Central Bank	Grantmaking and Giving Services
16	Office Administrative Services	Travel Arrangement and Reservation Services
17	Other Information Services	Junior Colleges
18	Data Processing, Hosting, and Related Services	Computer Systems Design and Related Services
19	Business Schools and Computer and Management Training	Management, Scientific, and Technical Consulting Services
20	Grantmaking and Giving Services	Other Information Services

Notes: This table lists the top 20 industries most exposed to AI from the original AIOE (Felten et al., 2021) and the top 20 industries most exposed to language modeling.