News Sentiment as a Dynamic Predictor of Job Automation Risk

Pierre Bouquet
Massachusetts Institute of Technology
Cambridge, MA, USA
pierre.bouquet.ch@gmail.com

Yossi Sheffi
Massachusetts Institute of Technology
Cambridge, MA, USA
sheffi@mit.edu

Amin Kaboli
École Polytechnique Fédérale de Lausanne
Lausanne, VD, CH
amin.kaboli@epfl.ch

Abstract

As artificial intelligence increasingly disrupts job and task structure, it is essential for companies and society, in general, to anticipate which tasks are at risk of automation and how these risks can guide workforce management strategies to proactively reskill employees, restructure roles, and optimize operations. To address these challenges, we introduce a machine learning pipeline that leverages news sentiment as a dynamic proxy for job automation risk assessment. By processing two million news articles, the model computes exposure scores at the task, job, and sector levels, enabling both historical trend analysis and real-time monitoring. Our findings demonstrate that these exposure scores align with prior studies that use rigorous, expert-driven methods. Through its dynamic evaluation, this approach models the impact of AI innovations and can help inform strategies for workforce transformation.

Keywords

Text Mining; News Sentiment; Predictive Modeling of Job Automation; Workforce AI Exposure

1 Introduction

Sentiment analysis is an application of natural language processing (NLP) to extract and quantify the underlying opinions and emotions expressed in text [87, 91]. The primary goal of sentiment analysis is to automatically identify and classify subjective information, typically categorizing it as positive, negative, or neutral [59, 23].

As Artificial Intelligence (AI) advances, the applications of sentiment analysis have expanded rapidly across diverse fields, including business [27], finance [25], politics [16], healthcare [14], and education [97]. For example, in the business domain, it is widely used to gauge consumer opinions [49], manage brand reputation [47], conduct market research [78], and improve customer service [50]. Hedge funds and other financial institutions use sentiment analysis to predict stock market trends and understand investor sentiment [85, 24, 25]. Political entities apply it to monitor public opinion [16] and predict election outcomes [77, 86]. Educational institutions use sentiment analysis to enhance teaching quality [97] and understand student learning capacity [79]. In healthcare, it helps to assess patient experiences [29] and track public health issues [14].

The ongoing breakthroughs in Deep Learning, particularly with Transformer-based Large Language Models (LLMs) [90] such as GPT [2], BERT [31], Gemini [44], and Llama [32], to name just a few, have significantly advanced the capabilities of sentiment analysis [65]. Day by day, these models become better at dealing with domain-specific context and the complexity of human language [95]. Coincidentally, the same technologies that could be summarized by the umbrella term "AI", also pose a risk of job automation since they are able to perform tasks traditionally attributed to human cognition [64, 63, 20]. Examples of such substitutions include the AI tools implemented by JPMorgan Chase to replace analyst work [41], Harvey AI's platform for legal work [57], and Duolingo applications in education [84].

For companies navigating the current wave of AI-driven disruption, it is important to monitor automation trends to anticipate changes in task and job composition [13]. Understanding which tasks are at risk of automation and how these risks evolve can guide workforce management strategies to proactively reskill employees, restructure roles, and optimize operations [83]. Existing rigorous and interpretable methodologies, such as Felten et al. [38] or Paolillo et al. [72], offer static snapshots of AI automation exposure and require highly laborious analysis. Such static approaches risk becoming outdated over time, as they often overestimate job displacement, underestimate productivity gains and the creation of new roles, and fail to account for ongoing innovations. Appendix A contains a detailed overview and discussion on the topic.

In contrast with existing methods, companies will benefit from a dynamic approach that provides real-time insight into the evolving risks of automation. In this paper, we show that the sentiment expressed in news articles regarding specific tasks or jobs can serve as a reliable predictor of automation risk. This intuition stems from several possible mechanisms. First, under the Rational Expectations Theory, corporate executives use all available information, including news, to make forecasts, resulting in decisions that align closely with those predictions [67, 45]. In this context, news sentiment helps shape the expectations and actions of executives as rational agents. Second, executives may be influenced by Herd Behavior, the tendency to mimic the actions of a larger group, often disregarding their own independent analysis [82]. News highlighting industry trends can amplify this effect as executives seek to align with peers. Lastly, this dynamic is further intensified by the Fear of Missing Out — a strong apprehension felt by individuals who believe that others are seizing rewarding opportunities from which they might be excluded [96, 6]. News emphasizing emerging AI trends can drive executives to act swiftly to avoid being left behind in adopting transformative technologies [94].

Our work demonstrates that the sentiment captured from news serves as a dynamic and reliable proxy for the risk of task automation. This is based on the assumption that news on AI-based innovations reflects both actual technological progress and the societal zeitgeist, which directly informs corporate decision-making regarding AI adoption. This perspective asserts that the trends and tendencies visible in news sentiment not only quantify the immediate risks of automation but also provide continuous, up-to-date insights into the impact of AI on the workforce.

To support our position, we demonstrate two key points. First, we show that with available technologies, it is possible to compute news sentiment in real-time or with reasonable latency, which we demonstrate by building a machine learning pipeline to generate exposure scores, as detailed in Section 2. Second, we demonstrate the reliability of news sentiment by showing that our exposure scores align with prior studies based on rigorous, expert-based methodologies, as discussed in Section 3. Section 4 provides broader insights and managerial implications for navigating workforce transformation in the age of AI.

2 Exposure Score

We developed a machine learning (ML) pipeline (Figure 1) that operationalizes the collection and analysis of sentiment data using the New York Times (NYT) and Agence France-Presse (AFP) as inputs. As explained below, using our ML pipeline, we design the AI Exposure Score at three levels: Task Exposure (TE) for 2,045 tasks, aggregated into Job Exposure (JE) for 873 jobs, and extended to Sector Exposure (SE) for 22 sectors.

2.1 Data and Model

We use three types of panel data spanning from 2019 to 2024: economic data, news articles, and jobs' structure (details are available in Appendix B). The economic data is readily usable and consists of occupational employment and wage statistics (OEWS) provided by the Bureau of Labor Statistics [88]. News articles are obtained from the The New York Times (NYT) and Agence France-Presse (AFP) outlets [3, 71]. They consist of unstructured text information. Finally, job structures are obtained from the ONET database, which provides detailed information regarding tasks, jobs, sectors, and their relationship [68].

We first parse the NYT and AFP data by removing incomplete entries. Then, we process each entry to add the topic and sentiment of each article. We use asymmetric Sentence-BERT (SBERT) embeddings to identify relevant articles, which capture the semantic essence of sentences by transforming natural language into a vector space [80]. We embed the content of news articles and compare their vectors against SBERT-vectorized prompts focused on AI automation. So, the cosine similarity score between these vectors quantifies the extent of their semantic alignment. The details on SBERT implementation and prompt engineering can be found in Appendix C.

We capture articles' sentiment using a Robustly Optimized BERT Pretraining Approach (RoBERTa) model [56]. The RoBERTa model, outputs a probability whether the text content is positive, neutral, or negative [58]. Positive sentiment means that the article posits that AI has the potential for increased performance, while negative sentiment discusses potential job displacements. Since the polar-opposite articles (strongly positive or negative) are likely to be too opinionated or politically motivated [5, 52], we build our sentiment score to favor more neutral articles. As a result, we have two quantitative features to describe each news article: its relevance to the topic of automation and its sentiment. Details about the sentiment score and RoBERTa are available in Appendix D.

We consider jobs as sets of tasks [11]. Although some jobs share tasks, the significance of a shared activity varies based on the time dedicated to it within each job. To model this, we use the frequency of each task for each job in the ONET database. We introduce the Annualized Task Time (ATT), which is an estimate of the minimum time commitment of a given task of a certain job at the time period for which the analysis is conducted. ATT serves an estimate of the number of a person's working hours that would be substituted if AI were to automate the task in question. Details about the ATT are available in Appendix E.

2.2 Task Exposure

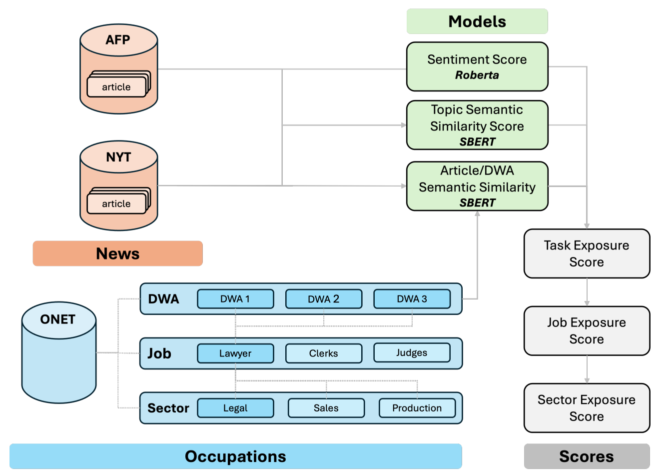

Figure 1: Machine Learning Pipeline

2.2.1 Tasks-Articles Relationship. The core of our AI exposure evaluation strategy is to build a continuously updated elemental time series metric. This section introduces our methodology to evaluate each task exposure (TE) using news articles. To evaluate the relationship between news articles and tasks, we calculate the cosine similarity score between the SBERT embeddings of the news article content (a) and the description of each task in ONET (i). We select articles that score above the 99th percentile for each task to ensure high relevance and specificity[1].

2.2.2 Task Exposure. TE at a time t for each task i is evaluated as a sum of the articles' sentiments weighted by their relevance to the topic of automation and the task (Equation (1)). This procedure ensures that more relevant articles have a more significant impact on the final score. As we consider news articles related to a task over a two-year time window, we normalize TEs to account for variations in the number of articles across different tasks, which makes the TE scores comparable between tasks.

$$TE_{t,i \in I_t} = \frac{\sum_{a \in A_{i,t}} s_a \cdot \omega_a \cdot \omega_{a,i}}{\sum_{a \in A_i} \omega_a \cdot \omega_{a,i}}, \quad (1)$$

Where $TE_{t,i \in I_t}$ is the AI Exposure of task i in the set I of tasks at time t; $A_{i,t}$ is the set of articles related to task i up to two years

[1]The percentile levels for topics and task cosine are calculated based on the complete dataset. Topic relevance is filtered to retain values at or above the 90th percentile, while task relevance is filtered to retain values at or above the 99th percentile.

prior to time t; s_a is the sentiment of article a; ω_a is the coefficient of semantic similarity between article a and the AI automation prompt; $\omega_{a,i}$ is the coefficient of semantic similarity between article a and task i.

2.3 Exposure of Jobs and Sectors

TE metrics provide insights into which tasks are most prone to automation. However, aggregated results are necessary to make statements regarding broader job trends. Therefore, we aggregate TEs for each time period by adding up tasks within each job, weighted by their ATT, and normalized to account for the disparity in the number of tasks per job. The results are job exposure (JE) metrics for 873 jobs (details available in Appendix F).

To evaluate sector exposure (SE), we evaluate 21 job families (referred to as sectors) by summing JE within each sector, weighted by the number of workers performing that job (based on the OEWS data) and normalized to account for the difference in the number of workers in each sector (details available in Appendix F).

Using news articles spanning 2017 to 2024, we apply our methodology to the entire dataset using a two-year rolling window. We incorporated the most up-to-date ONET data and the latest OEWS data available at the measurement time for each time window. For example, we use news data from January 1st, 2020, to December 31st, 2021, to compute TE on January 1st, 2022. We map tasks to jobs using ATT based on data from ONET version 26.1, released in November 2021, and generate JE scores. Finally, we group the jobs into sectors and weigh their importance using the May 2021 OEWS data. This procedure resulted in TE, JE, and SE metrics from 2019 to 2024, which we scaled between 0 and 1 to improve data interpretability [73]. The lowest TE, JE and SE are set to 0 while the largest values are set to 1. This procedure preserves relative difference in data points.

2.4 Most and Least Exposed Work

This section presents the most and least exposed sectors (SE), jobs (JE), and tasks (TE) during December 2023 and compares them to results in the literature.

2.4.1 Sectors' Exposure. The most exposed sectors to AI automation are predominantly white-collar, while the least exposed are blue-collar (See Table 1). Although this fact does not imply that blue-collar jobs are immune to automation, particularly by robotics, it indicates a lower immediate risk from AI automation than white-collar jobs.

Our findings align with several recent studies on AI's impact in various sectors. Felten et al. [38] similarly found sectors involving significant data processing and communication to be most exposed. For instance, the top two sectors (Legal and Education) are comprise predominantly desk jobs [7, 26]. Legal jobs are notably the most exposed due to the nature of their work, which involves processing large amounts of data to build cases. To wit, recent advancements in LLMs, such as OpenAI's GPT-4 model, scored 298/400 on the Uniform Bar Exam, placing it in the 90th percentile [2] of test takers, demonstrating the potential for automation in this field [57, 69].

The third most exposed sector, "Arts, Design, Entertainment, Sports, and Media," includes both field and desk jobs. This sector encompasses roles such as "News Analysts" and "Public Relations Specialists," which are highly exposed to AI automation [34], while "Craft Artists" within the same sector are less so. On the other end of the list, the three least exposed sectors to AI automation involve manual work that is difficult to automate by AI alone. However, as AI technology progresses new vulnerabilities may arise for manual work [61].

Additionally, the "Food Preparation and Serving Related" sector employs workers with low salaries and education levels, typically high school or equivalent. Investors are likely to favor inexpensive labor over investment in automation. We invite readers to interact with our data using our data visualization tool [2]

2.4.2 Job's Exposure. As SE scores are a weighted average of JE scores, understanding which jobs comprise these sectors is crucial for identifying those resilient to AI automation and those more exposed. Table 2 shows the three most and least exposed occupations out of 873 jobs evaluated.

The results are similar to the SE exposure, with certain white-collar professions particularly vulnerable to AI automation. The top three professions in Table 2 are especially vulnerable due to their heavy reliance on processing of data and executing routine cognitive tasks. For instance, JPMorgan has introduced a generative AI product that performs tasks typically undertaken by research analysts [41]. Also, AI technologies have advanced to the point where they can produce legal work comparable to that of a first-year associate, such as drafting memos and conducting preliminary research [69, 2].

Conversely, jobs requiring specialized manual skills or complex physical interactions have the lowest JE scores. This observation is similar to the conclusions of Kochhar [53], and Khogali and Mekid [51].

2.4.3 Tasks' exposure. Our analysis of JE metrics demonstrates that repetitive cognitive jobs are most exposed to AI automation, whereas manual specialized ones are more resilient. Table 3 lists the three most and least exposed tasks to AI automation.

Brynjolfsson et al. [20] developed a methodology to evaluate the suitability of tasks for machine learning and identified eight criteria to differentiate exposed from resilient ones. Tasks with clearly definable goals and metrics are more susceptible to automation. So are tasks that do not change rapidly over time and do not involve long chains of logic [20]. They also argue that tasks that do not require detailed explanations of the decision-making process and can tolerate some levels of error are more amenable to automation through AI.

The tasks with the highest TE scores, showed in Table 3, satisfy the criteria identified by Brynjolfsson et al. [20]. Tasks such as weighing parcels to determine shipping costs, determining geographic coordinates, and reading to students, involve simple, repetitive activities that do not require complex physical interaction. In fact, weighing parcels is already automated using a combination AI and robotics in many distribution centers, where material handling is performed entirely by automated systems [37, 81].

Several papers demonstrate that the use of AI and advanced geospatial technologies significantly reduces the need for human input in geographic data analysis [43, 62]. Finally, AI models can

[2] Visualization Platform Link

Table 1: SE Scaled Results on December 1st, 2023.

#	SOC Code	Sector Title	SE
1	23-0000	Legal	0.86
2	25-0000	Educational Instruction and Library	0.61
3	27-0000	Arts, Design, Entertainment, Sports, and Media	0.51
...	...	...	...
20	37-0000	Building and Grounds Cleaning and Maintenance	0.33
21	49-0000	Installation, Maintenance, and Repair	0.28
22	35-0000	Food Preparation and Serving Related	0.28

Table 2: JE Scaled Results on December 1st, 2023.

#	SOC Code	Job Title	JE
1	23-1012.00	Judicial Law Clerks	0.90
2	23-2093.00	Title Examiners, Abstractors, and Searchers	0.87
3	23-1022.00	Arbitrators, Mediators, and Conciliators	0.85
...	...	...	...
871	49-3041.00	Farm Equipment Mechanics and Service Technicians	0.42
872	35-2013.00	Cooks, Private Household	0.41
873	35-9021.00	Dishwashers	0.38

Table 3: TE Scaled Results on December 1st, 2023.

Row	SOC Code	Task Title	TE
1	4.A.1.b.3.I01.D13	Weigh parcels to determine shipping costs	0.96
2	4.A.2.a.4.I01.D02	Determine geographic coordinates	0.95
3	4.A.4.b.3.I02.D13	Read to students	0.91
...	...	...	...
2079	4.A.3.a.2.I19.D06	Remove parts or components from equipment	0.11
2080	4.A.3.a.2.I44.D02	Capture or kill animals	0.08
2081	4.A.3.a.1.I05.D05	Remove worn, damaged, or outdated materials from work areas	0.07

generate natural-sounding speech that has already proven useful in various applications, including virtual assistants, audiobooks, and educational tools [39, 89, 84].

Conversely, the least exposed tasks require manual dexterity and adaptability in changing environments, which makes them more challenging to automate. Tasks such as removing parts or components from equipment, capturing or killing animals, and removing worn, damaged, or outdated materials from work areas, involve manual dexterity and decision-making that are challenging for AI to replicate. The variability in equipment and parts and the need for precise physical manipulation make these tasks less susceptible to automation.

3 Is Exposure Score Reliable?

While many researchers provided the complete set of their JE scores, others only shared a sample of job exposure metrics. For the first category of research papers, which we refer to as "Complete," we perform a Pearson correlation test between our results and theirs. To ensure consistency, we use the JE obtained on January 1st of the year of publication, comparing metrics at the same point in time. For the second category, "Restricted," we compare their results to ours in terms of classification accuracy [76]. Finally, we provide an analysis of the robustness of the reliability of JE using different news datasets in Appendix 3.3.

3.1 The Complete publications

The Pearson correlations between our JE scores and the ones from several published papers are summarized in Table 4.

Table 4 shows that our JE metrics are mildly negatively correlated with studies on automation from rule-based software and computerization [42, 92]. The negative correlation suggests that while both AI and traditional software can automate certain tasks, they affect different types of jobs.

Our JE from AI is negatively correlated with studies on automation from robotics [72, 92]. This relation makes sense, as robots and AI do not impact the same labor force. AI is more likely to automate jobs performing cognitive repetitive tasks, while robots are more likely to automate tasks and occupations requiring physical skills and mobility [11, 7, 26, 19].

Table 4: Summary of Correlation Statistics from the Comparative Studies on JE

Study	Correlation Coeff.
Computerization and Rule-Based Software	
Probability of Computerization [42]	-0.29^{***}
Software Score [92]	-0.10^{**}
Robots	
Automation Risk Index [72]	-0.49^{***}
Robot Score [92]	-0.50^{***}
AI	
AI Occupational Exposure [38]	0.67^{***}
Suitability for Machine Learning [20]	0.12^{**}

Notes: **Correlation Coeff.** = Pearson correlation coefficient; * indicates statistical significance at $p < 0.05$; ** indicates statistical significance at $p < 0.01$; *** indicates statistical significance at $p < 0.001$.

Finally, there is a strong positive correlation between our JEs and those of Felten et al. [38], who also focused on AI. Our results and theirs align in identifying occupations exposed to AI automation, even though our methodologies are entirely different. In addition, our results show a positive correlation with the study by Brynjolfsson et al. [20], suggesting some agreement in identifying suitable jobs for machine learning, although the correlation is weaker.

3.2 The restricted publications

Some researchers shared only selected subsets of their results, sometimes highlighting only the most or least exposed occupations. To compare our results with those of these studies, we examined whether we identified the same most/least exposed occupations. For example, Kochhar [53] classified occupations into three categories: low, medium, and high exposure to AI. To obtain their classification, they divided their dataset into quantiles[3]. They shared 20 occupations sorted in alphabetical order within each quantile. To compare with their results, we divided our dataset using the same definition and compared the assessment of the 20 occupations from Kochhar [53] to ours. Table 5 shows that 58 percent of our results for JEs align with Kochhar [53]. Note that alignment is higher for high-exposure occupations and lower for low exposures. The high agreement for high-exposure occupations underscores the robustness of our model in identifying jobs critically exposed to AI. However, this JE alignment is limited since the article does not explain how the job samples were selected.

Eloundou et al. [34] do not provide detailed results on their tasks or jobs' exposure scores. However, they provide a list of 34 jobs without any task exposed to AI. Our methodology, though, does not classify jobs as having zero exposed tasks as it is a continuum from 0 to 1. Therefore, following Kochhar [53] we defined our bottom 25 percent of jobs as the least likely to be automated (218 jobs). The comparison resulted in a 79.4% match, meaning we similarly identified 27 out of the 34 jobs. This match suggests a strong alignment despite the broader scope of our study. While they focus specifically on large language models, our study encompasses AI at large.

[3]The 0th to 25th quantile represents occupations with the lowest exposure to AI, the 25th to 75th quantile represents medium exposure, and the 75th and above represent high exposure.

Table 5: JE Accuracy by Exposure Level with Kochhar [53]

	Low Ex.	Med. Ex.	High Ex.	W. Al.
Al. (%)	36.84	57.89	78.95	57.85
# Obs.	19	16	19	54

Notes: **Ex.** = Exposure; **W. Al.** = Weighted Alignment; **# Obs.** = Number of Observations. Some occupations (e.g., "Gambling services workers," "Teaching assistants," "Other Drafters") are not in the ONET database, thus no one-to-one association could be made.

3.3 Model Robustness

To assess the robustness of our framework to different news sources, we generated JE scores from 2019 to 2024 using data from the NYT and the AFP, both separately and in combination. Note that although different reporters write these news articles independently, they may cover similar events, meaning the NYT and AFP datasets are not entirely independent. It is also important to mention that both companies are considered high-quality news outlets, having won multiple reporting prizes [46, 54]. We analyze our framework's performance using the Pearson correlation alignment benchmark introduced in connection with the comparison to Complete publications. The results are summarized in Table 6.

Table 6 depicts similar performance across the two datasets. The JEs obtained using only AFP data provide equal or lower Pearson correlation coefficients than those obtained using NYT data. Additionally, the correlation is weak and statistically less significant, with one study using AFP data alone. This can be attributed to the difference in focus between AFP and the NYT, with the NYT being more US focused [4].

The combination of AFP and NYT datasets provides similar or better correlation coefficients (in amplitude) with existing research compared to each data source separately. In conclusion, our model is robust when used with high-quality news datasets, and larger datasets tend to perform better in our alignment benchmark test.

4 Insights and Managerial Implications

4.1 News Sentiment and Job Automation

Using empirical data, we reasonably showed that news sentiment serves as a dynamic and reliable proxy for assessing job automation risks. As suggested by Shook and Daugherty [83], managers, who are actively involving workers to shape the change and redesign their work and roles, may anticipate productivity gains of 20% or more over the next three years. In this regard, sentiment-based metrics offer managers a scalable tool for proactive task redesign, targeted reskilling, and scenario planning to address workforce transformation. Similarly, for policymakers, tractable metrics of AI exposure enable the identification of sector-level vulnerabilities, supporting the design of timely interventions to mitigate the societal impacts of automation [13].

Table 6: Summary of Correlation Statistics from the Robustness Studies on JE

Study	AFP	NYT	AFP & NYT
Computerization and Rule-Based Software			
Probability of Computerization [42]	-0.22^{***}	-0.29^{***}	-0.29^{***}
Software Score [92]	-0.09^{*}	-0.31^{***}	-0.10^{**}
Robots			
Automation Risk Index [72]	-0.35^{***}	-0.46^{***}	-0.49^{***}
Robotics Score [92]	-0.28^{***}	-0.52^{***}	-0.50^{***}
AI			
AI Occupational Exposure [38]	0.38^{***}	0.67^{***}	0.67^{***}
Suitability for Machine Learning [20]	0.10^{**}	0.15^{***}	0.12^{**}

Notes: The reported values are pearson correlation coefficient; * indicates statistical significance at $p < 0.05$; ** indicates statistical significance at $p < 0.01$; *** indicates statistical significance at $p < 0.001$.

4.2 Tasks Exposure as An Elemental Unit of Measurement

We also demonstrated that TE can be aggregated across various levels, from individual jobs to entire sectors. The flexibility resulting from building these exposure metrics based on an elemental unit (TE) means that one can easily apply this methodology at a company level. By mapping a company's labor force, TEs help identify tasks susceptible to automation. When combined with ATT and salary data, executives can simulate scenarios of workforce cost optimization and productivity enhancement. As AI adoption accelerates, TE can become a critical tool for AI governance, providing a framework for guiding regulations, processes, and technologies that ensure the alignment of AI deployment with an organization's strategic objectives [60]. Similarly, TE can be aggregated to model geographic AI exposure, as demonstrated by Felten et al. [38].

4.3 Managing Labor in the Age of AI

As of 2024, our analysis suggests that only three jobs—all in the legal sector—have all their tasks ranked in the 25 percent most exposed quartile. However, the legal sector contains seven occupations, suggesting that full automation is unlikely for all jobs, even in such highly exposed sectors. Combining this insight with the lower exposure of manual abilities highlights the varied nature of tasks exposed to AI. Thus, as the White House [93] argued, the primary concern surrounding AI is not the full automation of jobs but the disruption it brings across various tasks and sectors. Recent studies further show that AI can perform tasks typically carried out by entry-level workers, leading to increased productivity [41, 74, 28, 70]. However, there is concern that over-reliance on AI could reduce work quality [28]. This fact emphasizes the importance of collaboration between workers and AI to maintain high-quality output while improving efficiency [28, 30, 55]. Therefore, workers possibly have two primary options: they can either learn to collaborate with AI to enhance their productivity or transition into more manual roles, possibly in industries such as services, which are less exposed to AI [10, 55].

5 Limitations and Alternative Views

While our results demonstrate strong alignment with established methodologies, it is essential to discuss limitations and alternative perspectives.

First, the reliance on news sentiment may introduce biases inherent in editorial practices, such as sensationalism, partisanship, or partiality, which could skew the exposure scores [12, 17]. To address this, we propose expanding the dataset to include news outlets from diverse political backgrounds and applying sampling methodologies to balance the dataset [66]. Additionally, integrating non-news datasets, such as academic publications or patent data, could provide a more balanced and comprehensive view of automation trends [9, 92].

Second, sentiment-based dynamic approaches may be criticized for their susceptibility to transient news trends. Our framework addresses this concern by leveraging a two-year rolling time window, balancing the need to capture emerging trends with the goal of smoothing out short-term noise. Future work could explore the sensitivity of results to different time windows, incorporate exponential smoothing models to give greater weight to recent articles while maintaining the momentum of historical data, or hybridize the approach with expert assessments [18].

Lastly, the generalizability of our approach is constrained by its reliance on the ONET database, which reflects the structure of the U.S. labor force. To extend the applicability of our methodology globally, future research should adapt the framework to region-specific labor datasets, such as the European Skills, Competences, Qualifications and Occupations(ESCO) database [36].

6 Conclusion and Future Work

In this paper, we introduce an ML pipeline that leverages news sentiment as a dynamic and reliable predictor of job automation risks, addressing the limitations of static methodologies. Our pipeline quantifies each article's relevance to automation and specific tasks using SBERT and extracts sentiment polarity using RoBERTa. We apply this approach to a large corpus of news articles to generate features for our model, which computes AI exposure scores at the task, job, and sector levels. We validate our results through a comparative analysis, which demonstrates strong alignment with prior expert-driven studies. Finally, we discuss how organizations can leverage these insights for task redesign, targeted reskilling, and scenario planning.

In future work, we will explore the relationship between job exposure (JE) and workforce characteristics such as education level, wages, and task type (cognitive vs. manual). Additionally, we aim to extend our methodology beyond U.S. labor data by incorporating global workforce datasets such as ESCO, enabling cross-country comparisons of AI-driven labor transformations.

Acknowledgments

To Iona Clark, and Jaffar Namdar for their support and advice.

References

[1] Daron Acemoglu and Pascual Restrepo. 2020. Robots and jobs: evidence from us labor markets. *Journal of political economy*, 128, 6, 2188–2244.

[2] Josh Achiam et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

[3] AFP. 2024. Agence france-presse. https://www.afp.com. (2024).

[4] Charu C Aggarwal et al. 2015. *Data mining: the textbook*. Vol. 1. Springer.

[5] Mehwish Alam, Andreea Iana, Alexander Grote, Katharina Ludwig, Philipp Müller, and Heiko Paulheim. 2022. Towards analyzing the bias of news recommender systems using sentiment and stance detection. In *Companion Proceedings of the Web Conference 2022*, 448–457.

[6] Hephzibah Anderson. 2011. Never heard of fomo? you're so missing out. *The Guardian*.

[7] M Arntzi, T Gregoryi, and U Zierahni. 2016. The risk of automation for jobs in oecd countries. *Paris, Jun*.

[8] David Autor. 2022. The labor market impacts of technological change: From unbridled enthusiasm to qualified optimism to vast uncertainty. Tech. rep. National Bureau of Economic Research.

[9] David Autor, Caroline Chin, Anna Salomons, and Bryan Seegmiller. 2024. New frontiers: the origins and content of new work, 1940–2018. *The Quarterly Journal of Economics*, qjae008.

[10] David H Autor and David Dorn. 2013. The growth of low-skill service jobs and the polarization of the us labor market. *American economic review*, 103, 5, 1553–1597.

[11] David H Autor, Frank Levy, and Richard J Murnane. 2003. The skill content of recent technological change: an empirical exploration. *The Quarterly journal of economics*, 118, 4, 1279–1333.

[12] Kevin G Barnhurst. 2015. Contradictions in news epistemology: how modernism failed mainstream us journalism. *Media, Culture & Society*, 37, 8, 1244–1253.

[13] Niklas Berglind, Ankit Fadia, and Tom Isherwood. 2022. The potential value of ai—and how governments could look to capture it. *McKinsey*.

[14] Muzafar Bhat, Monisa Qadri, Majid Kundroo, Naffi Ahanger, Basant Agarwal, et al. 2020. Sentiment analysis of social media response on the covid19 outbreak. *Brain, behavior, and immunity*, 87, 136.

[15] Timothy Bresnahan. 2010. General purpose technologies. *Handbook of the Economics of Innovation*, 2, 761–791.

[16] Alexandros Britzolakis, Haridimos Kondylakis, and Nikolaos Papadakis. 2020. A review on lexicon-based and machine learning political sentiment analysis using tweets. *International Journal of Semantic Computing*, 14, 04, 517–563.

[17] Danielle K Brown, Summer Harlow, Víctor García-Perdomo, and Ramón Salaverría. 2018. A new sensation? an international exploration of sensationalism and social media recommendations in online news publications. *Journalism*, 19, 11, 1497–1516.

[18] Robert G Brown and Richard F Meyer. 1961. The fundamental theorem of exponential smoothing. *Operations Research*, 9, 5, 673–685.

[19] Erik Brynjolfsson and Tom Mitchell. 2017. What can machine learning do? workforce implications. *Science*, 358, 6370, 1530–1534.

[20] Erik Brynjolfsson, Tom Mitchell, and Daniel Rock. 2018. What can machines learn and what does it mean for occupations and the economy? In *AEA papers and proceedings*. Vol. 108. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 43–47.

[21] Bettina Buchel, Dario Floreano, et al. 2018. Tesla's problem: overestimating automation, underestimating humans. *Retrieved May*, 23, 2021.

[22] Pamela S Cain and Donald J Treiman. 1981. The dictionary of occupational titles as a source of occupational data. *American Sociological Review*, 253–278.

[23] Iti Chaturvedi, Edoardo Ragusa, Paolo Gastaldo, Rodolfo Zunino, and Erik Cambria. 2018. Bayesian network based extreme learning machine for subjectivity detection. *Journal of The Franklin Institute*, 355, 4, 1780–1797.

[24] Yong Chen, Bing Han, and Jing Pan. 2021. Sentiment trading and hedge fund returns. *The Journal of Finance*, 76, 4, 2001–2033.

[25] Zhuo Chen, Andrea Lu, and Xiaoquan Zhu. 2024. Investor sentiment and the pricing of macro risks for hedge funds. *Management Science*.

[26] Michael Chui, James Manyika, and Mehdi Miremadi. 2016. Where machines could replace humans-and where they can't (yet). *The McKinsey Quarterly*, 1–12.

[27] Jingfeng Cui, Zhaoxia Wang, Seng-Beng Ho, and Erik Cambria. 2023. Survey on sentiment analysis: evolution of research methods and topics. *Artificial Intelligence Review*, 56, 8, 8469–8510.

[28] Fabrizio Dell'Acqua, Edward McFowland III, Ethan R Mollick, Hila Lifshitz-Assaf, Katherine Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R Lakhani. 2023. Navigating the jagged technological frontier: field experimental evidence of the effects of ai on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, 24-013.

[29] Kerstin Denecke and Yihan Deng. 2015. Sentiment analysis in medical settings: new opportunities and challenges. *Artificial intelligence in medicine*, 64, 1, 17–27.

[30] Lydia DePillis and Steve Lohr. 2023. Tinkering with chatgpt, workers wonder: will this take my job? *International New York Times*, NA–NA.

[31] Jacob Devlin. 2018. Bert: pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

[32] Abhimanyu Dubey et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

[33] The Economist. 2024. Can artificial intelligence make health care more efficient? *The Economist*. Retrieved Aug. 10, 2024 from.

[34] Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. 2023. Gpts are gpts: an early look at the labor market impact potential of large language models. *arXiv preprint arXiv:2303.10130*.

[35] Brynjolfsson Erik and Mcafee Andrew. 2017. The business of artificial intelligence: what it can—and cannot—do for your organization. *Harvard Business Review Digital Articles*, 7, 3–11.

[36] European Commission. 2018. European skills, competences, qualifications and occupations (esco) (v.1.0.3). (2018).

[37] Stefan Fedtke and Nils Boysen. 2017. Layout planning of sortation conveyors in parcel distribution centers. *Transportation Science*, 51, 1, 3–18.

[38] Edward Felten, Manav Raj, and Robert Seamans. 2021. Occupational, industry, and geographic exposure to artificial intelligence: a novel dataset and its potential uses. *Strategic Management Journal*, 42, 12, 2195–2217.

[39] Tira Nur Fitria. 2023. Using naturalreader: a free text-to-speech online with ai-powered voices in teaching listening toefl. *ELTALL: English Language Teaching, Applied Linguistic and Literature*, 4, 2, 1–17.

[40] Martin Ford. 2015. The rise of the robots: technology and the threat of mass unemployment. *International Journal of HRD Practice Policy and Research*, 111.

[41] Joshua Franklin and Stephen Morris. 2024. JPMorgan pitches in-house chatbot as AI-based research analyst. *Financial Times*. Retrieved July 31, 2024 from.

[42] Carl Benedikt Frey and Michael A Osborne. 2017. The future of employment: how susceptible are jobs to computerisation? *Technological forecasting and social change*, 114, 254–280.

[43] Song Gao. 2020. A review of recent researches and reflections on geospatial artificial intelligence. *Geomatics and Information Science of Wuhan University*, 45, 12, 1865–1874.

[44] Google Research. 2023. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*. Accessed: 2025-01-01. https://arxiv.org/abs/2312.11805.

[45] Ilan Guttman, Ohad Kadan, and Eugene Kandel. 2006. A rational expectations theory of kinks in financial reporting. *The Accounting Review*, 81, 4, 811–848.

[46] John Hohenberg. 1959. *The Pulitzer Prize Story*. Columbia University Press New York.

[47] Akhilesh Ingole, Prathamesh Khude, Sanket Kittad, Vishakha Parmar, and Archana Ghotkar. 2024. Competitive sentiment analysis for brand reputation monitoring. In *2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE)*. IEEE, 1–7.

[48] International Federation of Robotics. 2023. *World Robotics 2023*. VDMA Verlag GmbH, Frankfurt am Main.

[49] Praphula Kumar Jain, Rajendra Pamula, and Gautam Srivastava. 2021. A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. *Computer science review*, 41, 100413.

[50] Daekook Kang and Yongtae Park. 2014. Review-based measurement of customer satisfaction in mobile service: sentiment analysis and vikor approach. *Expert Systems with Applications*, 41, 4, 1041–1050.

[51] Hisham O Khogali and Samir Mekid. 2023. The blended future of automation and ai: examining some long-term societal and ethical impact features. *Technology in Society*, 73, 102232.

[52] Brian Knutson, Tiffany W Hsu, Michael Ko, and Jeanne L Tsai. 2024. News source bias and sentiment on social media. *PloS one*, 19, 10, e0305148.

[53] Rakesh Kochhar. 2023. Which us workers are more exposed to ai on their jobs?

[54] Éric Lagneau. 2010. *Objectivity over the wire : the production of journalistic facts by the Agence France-Presse*. These de doctorat. Paris, Institut d'études politiques.
[55] Karim Lakhani. 2023. Ai won't replace humans–but humans with ai will replace humans without ai. *Harvard business review*.
[56] Yinhan Liu. 2019. Roberta: a robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 364.
[57] Steve Lohr. 2023. How Microsoft's Legal Legacy Shapes the Antitrust Case Against Google. *The New York Times*.
[58] Daniel Loureiro, Francesco Barbieri, Leonardo Neves, Luis Espinosa Anke, and Jose Camacho-Collados. 2022. Timelms: diachronic language models from twitter. *arXiv preprint arXiv:2202.03829*.
[59] Yukun Ma, Haiyun Peng, and Erik Cambria. 2018. Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive lstm. In *Proceedings of the AAAI conference on artificial intelligence* number 1. Vol. 32.
[60] Matti Mäntymäki, Matti Minkkinen, Teemu Birkstedt, and Mika Viljanen. 2022. Defining organizational ai governance. *AI and Ethics*, 2, 4, 603–609.
[61] James Manyika and Kevin Sneader. 2018. Ai, automation, and the future of work: ten things to solve for.
[62] Bruno Martins, Dalton Lunga, Song Gao, Shawn Newsam, Lexie Yang, Xueqing Deng, and Gengchen Mai. 2023. Report of the 5th acm sigspatial international workshop on ai for geographic knowledge discovery (geoai 2022). *SIGSPATIAL Special*, 14, 1, 23–25.
[63] John McCarthy et al. 2007. What is artificial intelligence.
[64] John McCarthy, Marvin L Minsky, Nathaniel Rochester, and Claude E Shannon. 2006. A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, 27, 4, 12–12.
[65] Md Saef Ullah Miah, Md Mohsin Kabir, Talha Bin Sarwar, Mejdl Safran, Sultan Alfarhood, and MF Mridha. 2024. A multimodal approach to cross-lingual sentiment analysis with ensemble of transformer and llm. *Scientific Reports*, 14, 1, 9603.
[66] Asmaa Mountassir, Houda Benbrahim, and Ilham Berrada. 2012. An empirical study to address the problem of unbalanced data sets in sentiment classification. In *2012 IEEE international conference on systems, man, and cybernetics (SMC)*. IEEE, 3298–3303.
[67] John F Muth. 1961. Rational expectations and the theory of price movements. *Econometrica: journal of the Econometric Society*, 315–335.
[68] National Center for O*NET Development. 2025. O*NET OnLine. (2025). https://www.onetonline.org.
[69] Jeff Neal. 2024. Harvard law expert explains how ai may transform the legal profession in 2024. *Harvard Law School*.
[70] Shakked Noy and Whitney Zhang. 2023. Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381, 6654, 187–192.
[71] NYT. 2024. The New York Times. https://www.nytimes.com. (2024).
[72] Antonio Paolillo, Fabrizio Colella, Nicola Nosengo, Fabrizio Schiano, William Stewart, Davide Zambrano, Isabelle Chappuis, Rafael Lalive, and Dario Floreano. 2022. How to compete with robots by assessing job automation risks and resilient alternatives. *Science robotics*, 7, 65, eabg5561.
[73] Fabian Pedregosa et al. 2011. Scikit-learn: machine learning in python. *the Journal of machine Learning research*, 12, 2825–2830.
[74] Sida Peng, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer. 2023. The impact of ai on developer productivity: evidence from github copilot. *arXiv preprint arXiv:2302.06590*.
[75] Peter Eckersley, Yomna Nasser, Yann Bayle, Owains Evans, Gennie Gebhart, and Dustin Schwenk. 2017. EFF AI Progress Measurement Project. (2017). Retrieved Apr. 16, 2024 from.
[76] David MW Powers. 2020. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.
[77] Miftahul Qorib, Rahel S Gizaw, and Junwhan Kim. 2023. Impact of sentiment analysis for the 2020 us presidential election on social media data. In *Proceedings of the 2023 8th International Conference on Machine Learning Technologies*, 28–34.
[78] Meena Rambocas and Barney G Pacheco. 2018. Online sentiment analysis in marketing research: a review. *Journal of Research in Interactive Marketing*, 12, 2, 146–163.
[79] Sujata Rani and Parteek Kumar. 2017. A sentiment analysis system to improve teaching and learning. *Computer*, 50, 5, 36–43.
[80] N Reimers. 2019. Sentence-bert: sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.
[81] Miguel Rodríguez-García, Iria González-Romero, Ángel Ortiz-Bas, and José Carlos Prado-Prado. 2024. E-fulfillment cost management in omnichannel retailing: an exploratory study. *Computers in Industry*, 159, 104094.
[82] Robert J Shiller. 1995. Conversation, information, and herd behavior. *The American economic review*, 85, 2, 181–185.
[83] Ellyn Shook and Paul Daugherty. 2024. Work, workforce, workers: reinvented in the age of generative ai. (2024).
[84] Natasha Singer. 2024. Will Chatbots Teach Your Children? *The New York Times*.
[85] David M Smith, Na Wang, Ying Wang, and Edward J Zychowicz. 2016. Sentiment and the effectiveness of technical analysis: evidence from the hedge fund industry. *Journal of Financial and Quantitative Analysis*, 51, 6, 1991–2013.
[86] Marty Swant. 2024. How AI shaped the 2024 election: from ad strategy to voter sentiment analysis. *Digiday*.
[87] Peter D Turney. 2002. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. *arXiv preprint cs/0212032*.
[88] U.S. Bureau of Labor Statistics. 2023. Employment by major industry sector. (2023).
[89] Aaron Van Den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, Koray Kavukcuoglu, et al. 2016. Wavenet: a generative model for raw audio. *arXiv preprint arXiv:1609.03499*, 12.
[90] A Vaswani. 2017. Attention is all you need. *Advances in Neural Information Processing Systems*.
[91] Mayur Wankhade, Annavarapu Chandra Sekhara Rao, and Chaitanya Kulkarni. 2022. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55, 7, 5731–5780.
[92] Michael Webb. 2019. The impact of artificial intelligence on the labor market. *Available at SSRN 3482150*.
[93] White House. 2022. The impact of artificial intelligence on the future of workforces in the european union and the united states of america. (2022).
[94] Xiaodong Yang, Bing Song, Liang Chen, Shirley S Ho, and Jin Sun. 2025. Technological optimism surpasses fear of missing out: a multigroup analysis of presumed media influence on generative ai technology adoption across varying levels of technological optimism. *Computers in Human Behavior*, 162, 108466.
[95] Wenxuan Zhang, Yue Deng, Bing Liu, Sinno Jialin Pan, and Lidong Bing. 2023. Sentiment analysis in the era of large language models: a reality check. *arXiv preprint arXiv:2305.15005*.
[96] Zhuofan Zhang, Fernando R Jiménez, and John E Cicala. 2020. Fear of missing out scale: a self-concept perspective. *Psychology & Marketing*, 37, 11, 1619–1634.
[97] Jin Zhou and Jun-min Ye. 2023. Sentiment analysis in education research: a review of journal publications. *Interactive learning environments*, 31, 3, 1252–1264.

A Evaluating Labor Automation: State of the Art

To review the state of the art in evaluating jobs' exposure to automation we first describe past efforts to evaluate the exposure from rule-based software and robotics. Next, we focus on the potential impacts of AI on tasks and jobs.
Key aspects of the methodologies and models reviewed include:

- Automation Mechanism: The type of automation technology.
- Measured Parameters: Each approach assesses automation's impact on specific parameters, broadly categorized into two groups: (1) Skills and abilities, and (2) Tasks.
- Input Data: The datasets used to evaluate AI exposure.
- Metrics: The specific measures of the potential impact of automation on jobs.
- Data Output: The format and type of the output.

A.1 Rule-Based Software and Robotic

Over the past three decades, the automation of repetitive information-processing tasks through rule-based software has significantly impacted the labor market, particularly in middle-wage occupations [11, 15]. Rule-based software systems are computer programs that implement manually specified decision rules [92]. Unlike many AI applications, a computer program requires the programmer to anticipate every contingency and code the necessary steps to complete tasks. Examples of typical applications include word processing, spreadsheet software, web browsers, and business applications like enterprise resource planning systems and help desk ticketing systems. Following Webb [92], who developed a model to assign a

Automation Mechanism	Authors	Measured Parameter	Input Data	Metric	Output Data
Computerization	Frey and Osborne [42]	Custom features	• ONET • Expert evaluation of 70 occupations	Probability of computerization	Snapshot
Computerization	Autor et al. [11]	Tasks	• IPUMS Census from 1960 to 1998 • Census local labor data	• Routine / nonroutine and manual/ cognitive TE • Impact of robots on wages	Economic insights
Robot	Acemoglu and Restrepo [1]	Task	• Industry exposure to robots (International Federation of Robotics, 2023)	• Impact of robots on employment	Economic insights
Robot	Paolillo et al. [72]	Abilities and Skill	• European H2020 Robotics Multi-Annual Roadmap (SPARC, 2016) • ONET	Average Retraining Effort	Snapshot
AI	Brynjolfsson et al. [20]	Tasks	• CrowdFlower evaluation of DWAs' automatability • ONET	JE	Snapshot
AI	Felten et al. [38]	Abilities	• AI Progress Measurement project by the EEF (Peter Eckersley et al., 2017) • ONET	Geographic Exposure	Snapshot
AI	Eloundou et al. [34]	Tasks	• Chat GPT4 evaluation of tasks' automatability • ONET	TE; JE	Snapshot
AI	Kochhar [53]	Tasks	• Survey data • ONET • Current population survey	Aggregated TE; JE; SE	Snapshot
Rule-based software; AI	Webb [92]	Tasks	• IPUMS 2010 census • ONET • Patents dataset	JE to software JE to AI	Snapshot
AI	**This paper**	Tasks	• News datasets • ONET	TE; JE; SE Annualized Task Time	5Y time series 1W timestep

Table 7: Summary of the Literature on the Impact of Several Automation Mechanisms on Workers

specific level of automation risk to each task, we define as "task exposure" (TE) metric in this study. These TE measures can be aggregated into "job exposure" (JE) metrics. Webb [92] argues that the impact of rule-based software varies by earning level, with middle-wage occupations being the most exposed. The declining demand for middle-wage occupations, coupled with increasing demand for both high- and low-wage positions, leads to job polarization: a phenomenon where wages and employment opportunities for middle-class workers shrink relative to those at the high and low ends of the wage spectrum [10, 92]. Webb [92] observed that as JE for certain jobs increased from the 25th to the 75th percentile, there were noticeable declines in both employment shares within industries (7-11%) and wages (2-6%). Additionally, he found a gender disparity in JE, with men more affected than women, likely due to the historical concentration of women in roles requiring complex interpersonal interactions.

Modern robots also use software to perform tasks. Robots are defined as automatically controlled, reprogrammable, and multipurpose machines [48]. Both Autor et al. [11] and Frey and Osborne [42] used the term 'computerization' to refer to the application of rule-based software automation to both workers and robots. Computerization can replace workers in carrying out routine tasks easily codified into programmed rules while supporting workers in performing non-routine tasks. With the significant decrease in computer costs over time, these dynamics (replacement and enhancement) have increased the demand for workers specializing in non-routine tasks, often those with a college education [11].

Using an expanded definition of computerization, which includes specialized machine-learning models to broaden the applicability of computerization to certain non-routine tasks, Frey and Osborne [42] reinforced that computerization poses a risk to low-skill, low-wage occupations. However, they still found specific hurdles to replacing higher-level tasks requiring sophisticated perception, fine motor skills, creativity, and interpersonal interaction.

As robots have become a transformative force in contemporary industries [40], certain researchers have narrowed their JE focus to only robotic automation. Webb [92] found that routine tasks with high levels of predictability are most at risk from robotic automation, especially in the manufacturing industry. The manufacturing sector, specifically the automotive industry, which utilizes 38% of industrial robots, is at the forefront of this transformation [1]. Acemoglu and Restrepo [1] found that in the United States, industrial robots have already had an impact on employment and wages, with each robot introduced per 1,000 workers decreasing the employment-to-population ratio by approximately 0.2 percent and average wages by 0.42 percent. For the most part, robotic automation challenged job security and wage levels for lower and middle-income workers [1].

Paolillo et al. [72] developed a JE score based on human/robot ability matching and derived an index called the "average retraining effort" designed to guide displaced workers toward the closest comparable occupations, least likely to be automated. It is important to highlight that current robots cannot automate entire manufacturing processes, as illustrated by Tesla's attempt to fully automate Model 3 production, which resulted in notable delays and production issues. Empirical evidence demonstrates that humans remain necessary, mainly due to their ability to adapt to unforeseen changes [21].

A.2 Automation through Artificial Intelligence

AI is commonly defined as the capability of a machine to imitate intelligent human behavior [63]. This definition typically refers to General AI, which implies that machines possess the ability to perform any intellectual task that a human can. However, such an advanced level of AI does not yet exist. This study focuses on machine learning algorithms, as defined in the introduction. Before we review the extant literature, we note that the impact of any technological evolution tends to take place gradually, influencing specific tasks rather than entire occupations [33]. AI is no exception, and its impact is multifaceted, encompassing both augmentation and replacement of various tasks [8].

The impact of AI on the workforce is likely to differ significantly from that of rule-based software and robotics [35]. In contrast with rule-based software, AI could reduce wage polarization across the labor force, except within the top earnings bracket, where it might increase inequality [92]. AI's advanced capabilities in pattern recognition, decision-making, and complex analysis can expose more skilled tasks and jobs to automation than earlier technologies. However, despite AI's potential to influence a wide range of tasks, Webb [92] argues that its capacity to fully automate jobs remains limited.

A specific type of ML architecture, large language models (LLMs), is trained on extensive datasets, where the input is a sequence of tokens (e.g., words), and the output is the next token in the sequence (e.g., next word in a sentence) [90]. Focusing on LLMs, Eloundou et al. [34] assessed TE by examining whether access to an LLM-powered system could reduce the time required for a human to perform a task. They found that approximately 80% of the U.S. workforce could see at least 10% of their tasks affected by LLMs, with about 19% potentially impacted by automation of at least 50% of their tasks. This study reinforces the findings of Erik and Andrew [35] that AI's impact is broad but unlikely to automate jobs fully. Like Webb [92], Eloundou et al. [34] suggest that many higher-income occupations face significant exposure to AI automation.

Unlike Brynjolfsson et al. [20], Webb [92], and Eloundou et al. [34], who defined TE as the base unit of AI exposure, Felten et al. [38] conceptualized jobs as aggregates of skills and abilities. They derived skills and abilities exposure from expert reports and developed a JE metric [75]. Additionally, they introduced an AI Industry Exposure metric called "sector exposure" (SE). Finally, they constructed an AI Geographic Exposure metric by aggregating SE data within a geographic area using employment statistics. Although Felten et al. [38] did not propose a method for updating JE, they emphasized that regularly updated employment data allows for ongoing monitoring of changes in exposure.

In conclusion, the most notable gap in the extant literature is that current studies offer only snapshots of existing exposure, without accounting for the rapid pace of ongoing innovations. This limitation is critical, due to the speed of improvements in AI technologies and the resulting labor disruptions. Therefore, there is an urgent need for models that can continuously assess the impact of AI on the workforce.

B Data Description

The datasets used to evaluate AI exposure scores consist of the Occupational Information Network (ONET) database [68], news articles from The New York Times [71] and Agence France Presse [3] for insights into public discourse on automation, and employment data from the U.S. Bureau of Labor Statistics [88].

B.1 Occupational Information Network - ONET

The ONET, is a database meant to serve as a source of occupational information [88]. Its goal is to provide workers, employers, and instructional designers with a shared language when it comes to occupations and skills. Workers can use it to explore different career options based on their skills and abilities. Employers identify skills necessary for their operations and thus, improve their efficiency of recruitment and training. Finally, educational planners use ONET to design training programs for the skills demanded in the workplace.

The concept behind the ONET was first introduced in 1938 in the Dictionary of Occupational Titles (DOT) [22]. It was published during an industrial economy, focused on blue-collar jobs. As the economy shifted away from heavy industry toward information and services, the DOT lost its usefulness.

It was later updated into the ONET database, with the first version released in December 1997, it was led by the U.S. Department of Labor's Employment and Training Administration which managed a team of public and private sector organizations. It included the data from the DOT as well as new data obtained from employers and workers. One of the biggest improvements is the focus on transferable skills making it easy to group jobs into related clusters and explore career paths across clusters. The dataset also identifies occupations by work activities allowing one to also group jobs based on their activities.

Today, the ONET provides comprehensive information about tasks, occupations, and industry sectors, utilizing the Standard Occupational Classification. Our study uses data from version 23.1 (November 2018) to version 28.3 (May 2024). The latest versions includes approximately 2,000 tasks mapped to 875 individual occupations, organized into 20 occupational families.

B.2 News Data

We used news data from two distinct sources, the New York Times and the AFP. The news coverage spun from 2017 to 2024.

B.3 Employment and Wage Data

We also utilized employment and wage data from the BLS, specifically the Occupational Employment and Wage Statistics (OEWS) for the period from 2019 to 2023 [88]. The OEWS program, conducted cooperatively by the BLS and State Workforce Agencies, surveys approximately 180,000 to 185,000 establishments every six months. It provides detailed data on jobs and wages across various industries and regions in the United States.

C Topic Filtering

The topic-filtering process assesses the relevance of articles to automation. After evaluating the relevance, we either remove non-relevant articles or retain pertinent ones. We employ an asymmetric Sentence-BERT (SBERT) model to analyze the article content. SBERT is a modification of the Bidirectional Encoder Representations from Transformers (BERT) architecture, designed to generate vector representations for sentences and longer texts, facilitating efficient and accurate semantic similarity comparisons. This comparison is conducted using cosine similarity, where a cosine value of 1 indicates strong alignment between the two vectors, and 0 signifies no relation. Asymmetric SBERT refers to the specific application of SBERT to tasks where the input texts differ significantly in length. In our case, one input (the automation topic prompt) is much shorter and simpler, than the news article content. We use the msmarco-bert-base-dot-v5 model, which maps sentences and paragraphs to a 768-dimensional dense vector space [80]. Our input structure is as follows:

- Prompt 1 - Short: "This paragraph discusses ________ and its relationship to a task or job."
  The blank is filled with terms related to automation, such as: ['automation', 'computerization', 'industrialization', 'artificial intelligence', 'expert systems', 'machine learning', 'neural networks'].Model Robst
- Prompt 2 - Long:
  - The content of the NYT articles consists of [headline] + [abstract] + [keywords].
  - The content of the AFP articles consists of [title] + [summary].

We compute the cosine similarity between each prompt and each article, selecting the maximum score as the automation relevance score for that article.

Finally, we filter out the least relevant articles using a quantile threshold, as detailed in the main body of the paper.

D Sentiment Model

The sentiment analysis of articles is performed using a Robustly Optimized BERT Pretraining Approach (RoBERTa) model [56]. This model outputs a probability distribution over three sentiment categories: positive, neutral, and negative [58]. To emphasize neutrality and discount highly polarized content, we compute a neutral sentiment score s_a for each article, which favors articles with a relatively balanced tone. The sentiment score s_a is defined as follows:

$$s_{a\in A} = 1 - p_{pos,a} - p_{neg,a} = p_{neut,a} \tag{2}$$

Where:

- $s_{a\in A}$: Neutral sentiment score of article a.
- $p_{pos,a}$: Probability of article a sentiment being positive.
- $p_{neg,a}$: Probability of article a sentiment being negative.
- $p_{neut,a}$: Probability of article a sentiment being neutral.

This formulation ensures that articles with strongly positive or negative sentiments receive lower scores, whereas articles with neutral sentiments receive higher scores. This scoring approach helps to filter out potentially biased or opinionated articles.

E Annualized Task Time

The Annualized Task Time (ATT) metric estimates the minimum time commitment for each task i within a given occupation j at time t, based on the task frequency data from the ONET database. ONET classifies task frequency using a seven-category scale, ranging from "Hourly or more" to "Yearly or less." Each frequency category is assigned a corresponding value, reflecting the estimated annual frequency of the task within the occupation. Assuming each task takes an equal amount of time to be performed, this provides a annual task time equivalent.

To calculate ATT, we assign weights to each task frequency category based on the minimum estimated occurrences per year. Thus:

$$ATT_{i,j,t} = 1344\cdot\alpha_{i,j,t}+672\cdot\beta_{i,j,t}+336\cdot\gamma_{i,j,t}+48\cdot\delta_{i,j,t}+12\cdot\zeta_{i,j,t}+2\cdot\eta_{i,j,t}+\theta_{i,j,t} \tag{3}$$

Where:

- $\alpha_{i,j,t}$: Value of the "Hourly or more" frequency of task i for occupation j in the ONET dataset at time t.
- $\beta_{i,j,t}$: Value of the "Several times daily" frequency of task i for occupation j in the ONET dataset at time t.
- $\gamma_{i,j,t}$: Value of the "Daily" frequency of task i for occupation j in the ONET dataset at time t.
- $delta_{i,j,t}$: Value of the "More than weekly" frequency of task i for occupation j in the ONET dataset at time t.
- $\zeta_{i,j,t}$: Value of the "More than monthly" frequency of task i for occupation j in the ONET dataset at time t.
- $\eta_{i,j,t}$: Value of the "More than yearly" frequency of task i for occupation j in the ONET dataset at time t.
- $\theta_{i,j,t}$: Value of the "Yearly or less" frequency of task i for occupation j in the ONET dataset at time t.

The multipliers in this equation (e.g., 1344) represent the minimum number of times each task is performed annually, based on the frequency categories from ONET. For example, a task categorized as "Hourly or more" (assigned 1344) reflects a lower bound of performing that task 52 weeks per year, at least 5 days per week, and at least once per hour over 8 working hours per day.

F TE Composite Score

The calculations of Job Exposure (JE) and Sector Exposure (SE), are derived from Task Exposure (TE), as discussed in Section 2.

F.1 Job Exposure – JE

To quantify the AI exposure at the job level, we calculate the Job Exposure (JE) score by aggregating the TE of tasks associated with a specific occupation. The aggregation is weighted by the Annualized Task Time (ATT) of each task within that job. The JE score for a given job j at time t is computed as follows:

$$JE_{j,t} = \frac{\sum_{i\in I_{j,t}} ATT_{i,j,t}\cdot TE_{i,t}}{\sum_{i\in I_{j,t}} ATT_{i,j,t}} \tag{4}$$

Where:

- $JE_{j,t}$: AI exposure of occupation j at time t.
- $I_{j,t}$: Set of tasks i associated with occupation j at time t.
- $ATT_{i,j,t}$: ATT of task i for occupation j at time t.
- $TE_{i,t}$: AI exposure of tasks i at time t.

Interpretation of Equation 4:

- The numerator represents the weighted sum of AI exposure across all tasks in the job, where each task is weighted by its ATT.
- The denominator is the sum of the ATT for all tasks in the job, which normalizes the result, ensuring the JE score reflects the AI exposure per job and is not biased by the number of tasks.

This metric accounts for both the AI susceptibility of individual tasks and their relevance to the job. It provides a view of how exposed each job is to automation, based on the task composition and time dedicated to those tasks.

F.2 Sector Exposure – SE

To assess AI exposure at a higher level of aggregation, we extend the analysis from jobs to sectors, defined as job families in the ONET database. Sector Exposure (SE) provides a measure of the AI exposure across 21 job families by combining the JE scores of occupations within each sector. This aggregation is weighted by the number of workers in each occupation, using data from the OEWS.

$$SE_{s,t} = \frac{\sum_{j \in J_{s,t}} \omega_{j,t} \times JE_{t,j}}{\sum_{j \in J_{s,t}} \omega_{j,t}} \tag{5}$$

Where:

- $SE_{s,t}$: AI exposure of sector s at time t.
- $J_{s,t}$: Set of occupations j in sector s at time t.
- $\omega_{j,t}$: Number of workers performing occupation j at time t (based on OEWS data).
- $OE_{j,t}$: AI exposure of occupation j at time t.

Interpretation of Equation 5:

- **The numerator** is the sum of JE scores for all occupations in the sector, weighted by the number of workers in each occupation.
- **The denominator** normalizes this sum by the total number of workers in the sector, ensuring the SE reflects the average AI exposure per worker in the sector.

This metric accounts for both the AI susceptibility of individual occupations and the distribution of workers across the sector. It provides a comprehensive view of how exposed each sector is to automation, based on the composition of jobs within the sector and the number of workers performing those jobs.

G Source Code and Visualization Platform

The experiments were conducted in a reproducible manner, and the source code, as well as data, are available on the GitHub repository: **ANONYMIZED**.

In addition we provide a platform to access and visualise the data with minimal technical requirements, it is available at: Visualization Platform Link.