Assessing AI's Disruptive Potential: A Novel Measure of Occupational Exposure Using Press Release Data

Georgi Demirev
Sofia University

August 9, 2024

Abstract

Recent rapid advances in Artificial Intelligence (AI) have sparked significant policy and academic interest in identifying occupations most at risk of automation. Previous studies addressing this issue relied on expert judgment, benchmark algorithm performance and patent data to determine similarities between AI capabilities and human tasks. However, these sources may not accurately reflect real-world AI applications. This study employs a novel dataset of over 440,000 corporate press releases to analyze AI product launches and adoptions. Using advanced natural language processing (NLP) techniques and large language models (LLMs), I identify and categorize tens of thousands of specific AI capabilities. These capabilities are then mapped to job-specific skills from the European Standardized Classification of Occupations (ESCO) to calculate an "AI Product Exposure Score" for over 3,000 individual occupations. The findings reveal significant variations in AI exposure across job categories, with Information and Communication Technology (ICT) workers, managers, and business administration professionals being the most exposed. Notably, these results indicate that higher-skilled and higher-paid professions are more exposed to AI, contrasting with previous automation waves that disproportionately affected lower-skilled workers.

JEL Classification: O33, J24, J23, O31, J31, C55, C81
Keywords: AI, job automation, technological change, occupational exposure

1 Introduction

Recent advances in Artificial Intelligence (AI) have sparked significant debate about the economic impacts these technologies will have as they spread across various industries. A particularly contentious issue is the potential of AI to automate jobs and displace human labor. In this respect, AI is often compared to previous waves of automation, such as software tools, industrial robots, and even electricity and the steam engine.

A critical question in this context is to what extent current AI capabilities overlap with and potentially substitute for tasks performed by humans. We can then look at what tasks are usually performed by various occupations and derive a measure of "exposure to AI" at the occupation level. This task-based framework, which views the production process as a series of steps, some of which can be automated, is a common approach in theoretical models of job automation (Acemoglu and Autor, 2011). In that light, identifying which occupations are at high risk of being affected by AI is a natural first step in modelling and predicting the broader impacts of AI as it develops further.

Several influential studies have sought to address this question. For instance, Frey and Osborne (2017) combines expert judgments with a machine learning classifier to produce AI automation scores for tasks and occupations within the O*NET occupational dataset. Building on this, Felten et al. (2018) and Brynjolfsson et al. (2023) utilize objective measures of AI advancements to assess the automation susceptibility of various tasks. More recently, Webb (2022) applies text analytics to a large corpus of patent data to identify AI capabilities, while Eloundou et al. (2023) employs a state-of-the-art large language model to judge whether a given task can be performed by AI.

In this paper, I propose a novel approach to estimate occupation-level exposure to artificial intelligence using data from company press releases. The key insight is that while academic papers, AI benchmarks, and patents might tell us what is feasible in principle, data on actual product releases and adoptions offers a more accurate picture of how AI is currently being implemented in the economy.

To achieve this, I collect a novel dataset of free-text corporate press releases from PRNewswire, a major press release aggregator, covering the period of August 2023 to April 2024 (prn, 2024). The dataset includes over 440000 press releases covering various topic from business and science to sports and culture. To identify the press releases relevant to this article - i.e. those concerning new products using artificial intelligence - I first apply a series of filters and then employ a combination of Natural Language Processing (NLP) techniques, such as word embeddings and vector similarity. This approach identifies 8507 AI-relevant press releases.

Each press releases is then further processed to remove any filler text or irrelevant details. The summarized articles are analyzed by a highly-capable language model (GPT-4) which is tasked with extracting specific "AI capability strings" from the text. These capabilities are meant to summarize the specific tasks that each of the analyzed AI products help to perform. Most of the AI products in the sample have between 3 and 4 associated capabilities, for a total of about 29000 non-unique capability strings.

The identified AI capabilities are then matched to a list of over 10000 job-specific skills from the classification of European Skills, Competences, Qualifications and Occupations (ESCO, European Commission, Directorate-General for Employment, Social Affairs and Inclusion (2019)). The matching is performed by first embedding each capability string and job skill and calculating pairwise cosine similarity between the two. By considering the relevance of each skill to each occupation in the ESCO taxonomy, I derive an "AI Product Exposure Score" for over 3,000 individual occupations.

The exposure score identifies occupations that require working with and analyzing information, as well as communication-intensive occupations, as being most exposed to disruption by Artificial Intelligence. On the other hand, occupations requiring manual labor and physical object manipulation are least vulnerable to recent advances in AI. Furthermore, the more exposed occupations are also more likely to have higher wages and lower unemployment, indicating that firms are employing AI for the tasks for which labor is most scarce.

The contributions of this paper are threefold: first, it introduces a novel dataset that has not been previously used in this field. Second, it employs an innovative combination of NLP and LLM techniques to analyze this data, which could be applied to other datasets as well. Third, it presents findings at the level of ESCO occupations, making them compatible with the classification system used by the International Labor Organization, unlike prior research that relies on the US-specific O*NET classification.

The rest of the paper is organized as follows. Section two reviews the related literature. Section three details the methodology of the analysis. Section four presents the main results, including the occupation exposure score and supplementary findings such as correlations with occupation-level statistics and comparisons with prior studies. Section five offers a discussion and concludes the paper. All code used to produce this paper, as well as all resulting files are available online[1].

2 Literature Review

Previous episodes of rapid technological change have often been accompanied by equally dramatic shifts and adjustments in the labor market Acemoglu and Johnson (2023). As such, it is unsurprising that the issues of job automation and labor displacement have received significant academic interest from both theoretical and applied economists.

On the theoretical side, the task-based framework first proposed by Zeira (1998) has been a workhorse model used to conceptualize the effects of automation on economic production. In this framework, economic output is produced by aggregating different "tasks", which are themselves a product of some combination of labor and capital. Technological advances are in turn thought of as improvement of capital productivity in some subset of tasks. This leads to the employment of more capital relative to labor in these tasks (i.e. automation). Keeping the total number of tasks constant, ultimately this leads to the reduction of labor's share of output.

This framework has been further developed and refined by several works by Acemoglu and Restrepo, such as Acemoglu and Restrepo (2018), Acemoglu and Restrepo (2020), and Acemoglu and Restrepo (2019). An unifying finding of these studies is that technological improvement can affect total production in four different ways depending on how exactly it affects labor and capital. The first two are labor augmentation and capital deepening, i.e. an across-the-board improvement in either labor or capital productivity or both. This is analogous to Harrod-neutral or Solow-neutral technological change in standard growth theory and it's effect are thus the familiar increase in total output while (under standard assumptions) keeping the labor share constant.

More interestingly, technology may change the set of tasks that are produced by capital in equilibrium (i.e. automation) or it may lead to the creation of new tasks. Automation has straightforward labor displacement and labor share reduction effect. On the other hand the creation of new tasks always has a positive effect on the labor share of output Acemoglu and Restrepo (2018). Therefore the effect of any technological advancement (including AI) on labor demand is ambiguous and ultimately depends on how the technology impacts the economy at the task level.

Another feature of these models is that even if technology ends up acting as pure automation, it can still end up with no adverse effects on labour. This is because the resulting increase in productivity may be big enough to grow the entire economy in such a way that the overall increased demand for labor compensates for task-level displacement Acemoglu and Restrepo (2019). A famous case-study illustrating this effect concerns bank tellers and ATM machines and is presented in Bessen (2015). Even though ATMs were introduced to automate parts of the job of the tellers, the resulting cost savings were so great, that it allowed banks to expand and open more branches, with the overall result being the employment of more bank tellers compared to before the introduction of the ATM. This tension between displacement on one hand and complementarity plus productivity gains on the other, underscores the necessity of examining AI's granular impacts at the task and occupation levels.

In another influential theoretical work using a task-based approach, Aghion et al. (2019) investigate whether inherent complementarities in the production function might limit AI's economic impact due to Baumol's cost disease (Baumol, 1967). The authors argue that because the outputs of different tasks are gross complements in the final good production function, tasks automated by AI might follow the path of agricultural jobs, in the sense that as their productivity rapidly

[1] `https://github.com/demirev/ai-products/`

increases, the share of those tasks in the economy will dramatically decrease. This conclusion reinforces the need to understand which tasks and occupations may experience the most disruption due to AI.

Aghion et al. (2019) also explore the potential for AI to automate research labor in the ideas production function. In most endogenous growth models (e.g. Jones (2022) or Romer (1986)) total factor productivity growth is modelled as a separate production process ("idea-generating function") with its own inputs. The most important of these inputs is research labor, and the long-term growth rate is ultimately limited by the growth of the number of researchers. If AI is able to automate certain research tasks, it may be viewed as an effective increase in research labor, and thus have implications on long-term growth. Aghion et al. (2019) model the idea-production function itself as a combination of tasks, some of which can potentially be automated. This warrants a closer examination of AI's applicability to research-related tasks.

The theoretical arguments presented above underline the need for assessing the susceptibility of particular tasks and occupations to automation by AI. To address that need, several measures of occupational AI exposure have been proposed in the literature. Pioneering work in this direction includes Brynjolfsson et al. (2018) and Brynjolfsson and Mitchell (2017). These studies introduce the approach of evaluating occupations based on their constituent tasks and assessing each task's susceptibility to being performed by AI—a methodology broadly followed by subsequent literature and one that this paper also employs.

More specifically, Brynjolfsson and Mitchell (2017) develops a "Suitability for machine learning" rubric that aims to measure the degree to which each of the 2,069 direct work activities in the O*NET database is suitable for automation by AI. They do this by employing seven separate people via a crowd-sourcing platform and having them answer a set of 23 questions for each activity. Brynjolfsson et al. (2018) then uses this rubric at the occupational level to examine if some occupations are fully susceptible to AI automation. They find that most occupations consist of a mix of ML-suitable and unsuitable tasks, but also underscore the huge variability across occupations. Although Brynjolfsson et al. (2018) finds weak correlation between their suitability measure and wages, the authors note that firms are likely to employ the technology directionally and focus on the highest paid tasks among those that are suitable for automation with machine learning.

Frey and Osborne (2017) builds on this approach by developing a machine learning classifier to assess the AI exposure of individual occupations. They begin by constructing a hand-labeled dataset of 70 occupations rated by machine learning experts. Next, they employs occupational characteristics from the O*NET database (National Center for O*NET Development, 2024) as predictor variables and fit a Gaussian-process classifier to predict the exposure score based on the predictors. Finally, they feed data on 702 detailed occupations into this classifier to obtain estimates for AI susceptibility for all occupations. Interestingly, their approach identifies tasks requiring creativity ans social intelligence as bottlenecks for AI adoption, a result that may no longer hold as strongly with the advent of generative AI. This paper also finds that lower-paid and lower-educated workers may face higher probability of computerization - a result in contrast with the findings of some of the other cited studies.

To introduce an objective measure of AI's technical progress, Felten et al. (2018) utilizes AI progress metrics compiled by the Electronic Frontier Foundation (Eckersley and Nasser, 2017-). This dataset tracks AI capabilities across 16 categories, such as image recognition, question answering, and video games. The authors link these metrics to 52 O*NET occupational abilities using hand-crafted relevance weights. The weights are based on the expert judgement of PhD students in computer science. Next, using the slope of the progress for each metric, they identify which tasks are likely to be most affected by AI algorithms in the near future. Finally, by aggregating the tasks to the occupational level, Felten et al derive an "AI Occupational Exposure" measure.

In subsequent work Felten et al. (2023), the same authors augment their score to focus specifically on large language models. One of the interesting findings of this study, is that the AI exposure score exhibits statistically significant correlation with scheduled occupation definition changes in the O*NET database, showing that highly exposed occupations indeed exhibit some form of disproportionate change or disruption. Felten et al's data is publicly available and is used in this paper to validate its results.

Webb (2022) uses text analytics to analyze patent titles and descriptions, extracting AI capabilities and matching them to occupational abilities. The patent data is sourced from Google Patents Public Data, while the occupational abilities are the natural language description of tasks related to 964 occupations from O*NET. Webb filters patents that are related to AI and then extracts verb-noun pairs from both the patent titles and task descriptions. These verb-noun pairs are then processed and generalized using a hierarchical database of word meaning. The final occupational exposure measure calculated by Webb is based on the overlap in the verb-noun pairs between occupation task descriptions and patent titles.

Using this approach, Webb (2022) finds that AI is more likely to affect higher educated and higher paid workers, as well as occupations with a higher percentage of male workers or workers over 30. This study conceptually aligns with Webb's approach but uses product press releases instead of patent filings. While Webb relies on verb-noun pairs and hierarchical concept substitution, this paper utilizes vector similarity for a more nuanced measurement. Webb's scores are also available online and compared to the measure proposed here.

Another paper that shares methodological similarities with this study is Eloundou et al. (2023). The authors outsource the task of rating AI exposure to a Large Language Model (LLM). By crafting specific prompts for GPT-4 (OpenAI, 2024), they derive consistent scores across occupations. The definition of exposure in their paper requires that the time needed to complete a given task could potentially be reduced by 50% if an LLM is employed to complete it. They further breakdown their exposure metric in three categories: "no exposure" if no or minimal productivity gains can be expected, "direct exposure" if significant productivity improvements can be obtained by applying an LLM "as is", and "LLM+ exposure" if additional systems or products are required to achieve significant productivity gains. To demonstrate the robustness of their approach, the authors personally labeled a subset of the occupation-relevant tasks and compared the ratings to the ones obtained by GPT-4. This paper also uses GPT-4, but only as a final step in text processing, as outlined in the methodology section.

Overall, Eloundou et al. (2023) finds that between 15% and 56% of all tasks can potentially see significant productivity gains due to LLMs in the near future. The study also finds that higher-paid and higher-educated occupations tend to be more exposed to AI. The results of Eloundou et al are also freely available and used for the validation in the Results section.

3 Methodology

3.1 Data

This study utilizes three main data sources: the PRNewswire repository of press releases, used to extract AI product capabilities (prn, 2024); the European Skills, Competences, Qualifications and Occupations (ESCO) job taxonomy data, used to map these capabilities to occupational skills and capabilities (European Commission, Directorate-General for Employment, Social Affairs and Inclusion, 2019); and the Skills Intelligence Report from the European Centre for the Development of Vocational Training (CEDEFOP), used to examine job market and demographic statistics across occupations (European Centre for the Development of Vocational Training, 2024).

PRNewswire is a prominent online aggregator of corporate press releases, distributing and publishing them on its publicly accessible website. This website contains data for the last six months spread across 14 categories. For the purposes of this study, I collected the full text of all available press releases from August 8, 2023, to April 15, 2024, totalling 444,989 individual press releases.

The ESCO dataset contains a detailed multi-level hierarchical taxonomy of occupations. While the first four levels of classification align with the International Standard Classification of Occupations (ISCO) by the International Labour Organization (ILO), ESCO further subdivides the categories, resulting in five levels of occupational detail. For example, the occupation "bank teller" (ESCO code 4211.1) falls under "Tellers and money collectors" (ESCO code 421), within "Customer Service Clerks" (ESCO code 42), which is part of "Clerical Support Workers" (ESCO code 4). The version of the ESCO database used (v1.1.2) includes 3,006 occupations and 10,821 job-relevant skills.

The Skills Intelligence Report is an online publication compiled by CEDEFOP based on several primary data sources, such as the European Labor Force Survey (EU-LFS) and the European Survey on Income and Living Conditions (EU-SILC). It contains occupational-level data for all EU countries aggregated at the 2-digit ESCO level. Key data used from this report includes the number of workers per occupation per country, unemployment rates, and the percentage of women employed in each occupation. This dataset allows for linking the occupational exposure index developed below to key EU-level job market outcome statistics.

Additionally, data from the European Commission's Wage Determinants report (European Commission. Statistical Office of the European Union., 2021) is used to assign a wage premium to each occupation. This report contains an econometric model analyzing wage determinants, using 2-level ESCO code dummies. The coefficients on these dummies are interpreted as relative wage premiums by occupation.

3.2 Text Analysis

The text analysis of the press releases involves two main steps: 1) identifying press releases about new AI product launches or the adoption of AI-powered products, and 2) identifying the capabilities of these AI systems.

Firstly, since PRNewswire is a general-purpose publication, I apply a keyword filter to limit the number of documents for analysis. After standard preprocessing (lowercasing, removing stop words and punctuation, stemming), I check for the presence of phrases related to artificial intelligence (e.g., "AI," "machine learning," "gpt," "openai"). This filtering process identifies 69,661 potentially relevant articles. I denote this document corpus as P.

Next, I use a semantic similarity comparison to determine which articles pertain to AI product launches and introductions. Using an open-source NLP machine learning model (Honnibal et al., 2022), I perform text embedding on all product releases and benchmark phrases [2].

Formally, if $E()$ is the embedding function that acts on an individual word w to produce a K-dimensional embedding vector, the document-level embedding of each document $r \in P$ is denoted as $\bar{E}(r)$, whose the k-th element is given by:

$$\bar{E(r)}_k = \frac{\sum_{w \in r} E(w)_k}{\sum_{w \in r} 1} \tag{1}$$

i.e. an element-wise average of all individual word embeddings in the document.

[2] The particular model used is version 3.7.1 of the model en_core_web_lg of the open-source package SpaCy, which produces 300-dimensional embeddings.

I define several phrases indicative of AI product launches or adoptions[3] and derive their embeddings in an analogous fashion. I then compute the cosine similarity between the press release embeddings and benchmark phrase embeddings. Using the $\bar{E}()$ function defined above, the cosine similarity between any two documents (in this case a press release and a key phrase) is defined as:

$$S(i,j) = \frac{\bar{E}(i) \cdot \bar{E}(j)}{\|\bar{E}(i)\| \, \|\bar{E}(j)\|} \tag{2}$$

Having computed the similarity measure between all press releases in P and each key phrase, I filter out press releases with similarity scores below 0.7 across all key phrases, resulting in 8,507 relevant press releases. I denote this document corpus as R.

The next phase of the analysis involves identifying the specific capabilities of the AI tools mentioned in the press releases. I apply a similar semantic similarity exercise at the sentence level to identify relevant text segments. I.e. for each document $r \in R$ I retain only those sentences s for which $S(s,j) \geq 0.8$ for at least one key phrase j. This is needed to reduce the overall text size and hence the processing costs for the next step of the analysis.

Finally, I pass the condensed version of each press release (containing only the sentences scored as relevant) to a large language model (GPT-4). The model is prompted to read the text and output short, comma-separated strings detailing the AI system's capabilities (e.g., "reads and analyzes documents, issues treatment recommendations")[4]. This format mirrors the occupation capability strings in the ESCO database. The model outputs an empty string if the text is not relevant or if no specific capabilities are identified.

Formally, if $U(r)$ is the set of AI capability phrases, identified by the Large Language Model based on input document r, then the set of AI capability phrases is defined as:

$$C_{ai} = \bigcup_{r \in R} U(r) \tag{3}$$

This process results in 28,908 AI capability strings coming from 5,300 press releases.

3.3 Deriving Occupational Exposure Scores

The identified AI capabilities are compared to the 10,821 individual occupational skills in the ESCO database using cosine similarity between the mean word embeddings of each string. This generates nearly 313 million pairwise semantic similarity scores.

For each individual skill, I take the highest score from these comparisons, and assign it as that skill's exposure to AI product innovations. Formally, if $i \in C_{jobs}$ is an individual jobs skill string, and $j \in C_{ai}$ is an individual AI capability string, and $S(i,j)$ is the similarity between i and j as defined above, the skill-level similarity score $S^*(i)$ is given by:

$$S^*(i) = \max_{j \in C_{jobs}} S(i,j) \tag{4}$$

Finally, I aggregate these similarity scores up to the 5-digit occupation level. I do this by taking the list of skills associated with each occupation in the ESCO database and computing a weighted average of the individual skills' exposure scores. I weigh skills labeled as "essential" twice as heavily as skills labeled as "optional" [5]. Formally:

[3] See Appendix

[4] The prompt can be found in the Appendix

[5] This weighting is admittedly arbitrary, but the results do not meaningfully change if I employ an unweighted average instead

$$ES(o) = \frac{\sum_{i \in JS(o)} w_{o,i} S^*(i)}{\sum_{i \in JS(o)} w_{o,i}} \tag{5}$$

Where $ES(o)$ denotes the (non-standardized) occupational exposure score of occupation o, $JS(o)$ is the set of job skills relevant to occupation o, and $w_{o,i}$ is the importance weight of the skill for the occupation (either 1 for non-essential or 2 for essential skills).

Since this is essentially a dimensionless index, I standardize the result by subtracting the mean and dividing by the standard deviation. The resulting quantity $ES^*(o)$ is the **occupation-level exposure to AI product innovations score**, the principal output of this study.

When presenting the exposure scores for broader ESCO categories, I use a simple average. For example, the scores at the 3-digit occupational level are the average of the corresponding 4-digit occupations. I also apply a crosswalk file provided by ESCO to convert the scores to the United States Standard Occupational Classification for comparisons with prior work. All final scores are available through the open-source repository associated with this paper.

3.4 Limitations

Despite being the main contribution of this paper, the database of corporate press releases used to extract AI product capabilities also has several limitations that need to be acknowledged. First, it introduces a bias towards larger companies and more publicized AI implementations. Smaller companies or those that operate in sectors where publicity is not highly valued are likely to be underrepresented in the analyzed dataset. Furthermore, companies have an incentive to embellish and oversell their products in promotional materials, meaning that some of the identified capabilities may not reflect the real use cases of the products.

Additionally, due to limitations in data availability, the analysis covers press releases from August 2023 to April 2024 - a relatively short time frame. While this time window likely coincides with the early commercialization of current-generation Large Language Models, it may not be sufficient to capture longer-term trends in AI development and adoption. Moreover, by focusing on product launches and adoptions, the methodology may not fully capture AI's impact on internal business processes that aren't marketed as products.

This paper also uses two separate commercially available large language models - SpaCy for text embeddings and GPT-4 for the final step of extracting product capabilities. This means that the results are sensitive to the specific training and validation procedures used when developing these models, which are not always easy to interpret. Furthermore, the outputs of GPT-4 are stochastic in nature and can vary if the seed parameter or model version are changed.

Another important limitation is that the methodology doesn't distinguish between AI applications that augment human capabilities versus those that potentially replace human labor entirely. While such information might be present in the text of the press release, the proposed methodology is not designed to identify it. This nuance could be crucial in understanding the true impact of AI on different occupations.

Finally, the use of crosswalks to compare results with previous studies based on different occupational classifications may introduce some inaccuracies. These limitations should be considered when interpreting the results and could serve as starting points for future research and methodological refinements.

4 Results

4.1 Job Skills and AI Product Capabilities

The ESCO database groups skills in several hierarchical levels (similar to occupations). There are eight 1-digit groups, describing all job skills in broad terms, as well 66 2-digit groups and 234 3-digit groups, giving more detailed break-downs of job relevant skills.

The average of the individual skill exposure scores (defined as the highest cosine similarity between the skill string and the AI capability strings as per the previous section) for each 1-digit and 2-digit skill groups are presented in Figure 6 and Table 2 in the appendix.

The 1-digit skill group most overlapping with new product AI capabilities is "Working with computers", followed by management, information skills, and communication, collaboration and creativity. On the other hand handling, moving, and working with specialized equipment have relatively little similarity with the capabilities of AI products.

At the 2-digit skill group level, the most exposed skills include technical abilities such as programming, analyzing data, and managing information, as well as some people-related skills such as building teams, counselling, setting objectives. The broad skill "solving problems" is also towards the top in terms of AI exposure.

Curiously, some of the skills related to producing graphical content and written text, such as "writing and composing" and "creative artistic and visual materials" don't score particularly high on the exposure metric, despite the advances of generative AI in these area. This perhaps indicates that, impressive as those capabilities may be, companies haven't found ways to effectively commercialize them yet.

4.2 Occupational Exposure Scores

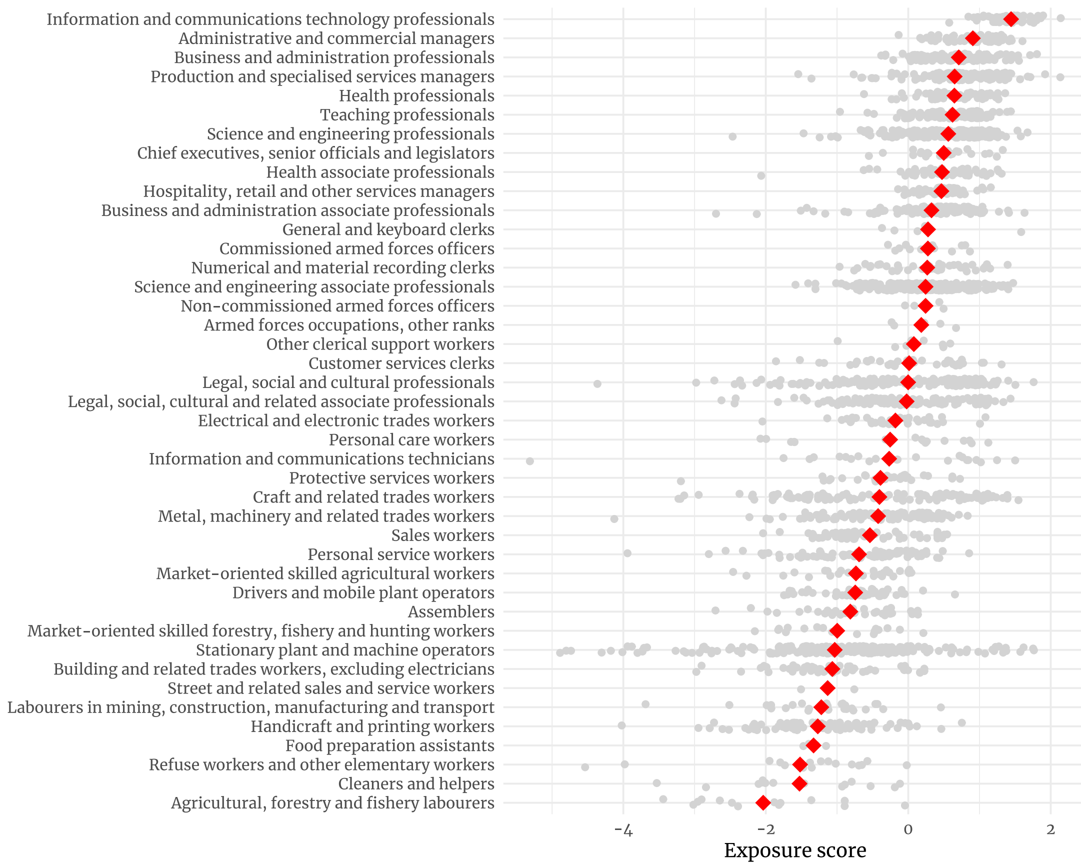
The main result of this study, the occupational exposure scores, are summarized in Figure 1. Each labeled row corresponds to an occupational group at the 2-digit level of the ISCO hierarchy. The gray dots represent individual (5-digit) occupations within each group, while the red rhomboids represent group means. The group means are also presented in tabular form in **??** in the appendix.

ICT professionals are the most exposed group with an exposure score of 1.44, followed by Administrative and Commercial Managers (0.907) and Business and Administration Professionals (0.709). Also notable near the top of the table are Health and Teaching Workers, Science and Engineering Professionals, and Senior Executives. At the opposite end, Agricultural, Forestry, and Fishery Laborers are the least exposed group, followed by Cleaners and Helpers, Refuse Workers, and Food Preparation Assistants.

The exposure scores can also be aggregated on the level 3 or 4 digit occupation groups. The corresponding tables are included in the online repository, while Section 4.2 present the 10 highest exposed 3 digit occupations. At this level of aggregation, the occupations that can expect the most AI disruption include software-related jobs such as developers and database technicians, as well as medical doctors and nurses, university educators, as well as various management-related occupations. A similar table is available in the Appendix for the lowest scoring 3-digit occupations (Appendix A.3). The least exposed professions include various jobs related to physical labor, such as manufacturing workers, craftspeople, machine operators, and cooks.

At the individual occupation level, the most exposed jobs are "Data Analyst", "Chief Data Analyst", "ICT Information and Knowledge Manager", "ICT Business Analyst", and "Software Analyst", suggesting that the engineers developing AI systems are first focusing on improving the efficiency of their own work. The full list of 3,007 scored occupations is available in the online repository of this paper.

Figure 1: Occupational Exposure to AI Product Innovation



A specific group of occupations and skills that warrants a closer look is research-related professions. As mentioned above, AI advances in research and knowledge production have the potential to bring long-lasting productivity and welfare gains (Aghion et al., 2019). ESCO provides a list of skills especially relevant for research work. According to the measure developed in this paper, those skills have a high susceptibility to AI disruption, with a mean score of 0.93 compared to 0.0 (by construction) for all other skills. Similarly, research-related professions score highly on the occupation exposure level, with Science and Engineering Professionals having a mean score of 0.562. This result is presented graphically in Figure 2, where the overall distribution of per-skill exposure scores are presented with the solid density line, while research skills are plotted individually as red dots along the x-axis.

Overall, the profile of the occupations most exposed to disruption from new AI products hints that whatever the labor market effects of AI may be, they will differ from previous waves of automation (e.g. industrial robots that targeted mostly manual and lower skilled labor, as studied by Acemoglu and Restrepo (2020). The jobs that are most likely to see any effects are concentrated in knowledge work (programming, data analysis, research), various forms management, as well as education and health care.

4.3 Comparison to Prior Work

To validate the results and compare them to prior work, the final result datasets of Felten et al. (2018), Eloundou et al. (2023), and Webb (2022) were downloaded and joined by occupation to

Table 1: Highest Exposure Score Occupations at the 3-digit ISCO Level

ISCO Code	Exposure Score	Occupation
251	1.46	Software and applications developers and analysts
133	1.41	Information and communications technology service managers
252	1.41	Database and network professionals
221	1.15	Medical doctors
351	0.948	Information and communications technology operations and user support technicians
222	0.946	Nursing and midwifery professionals
122	0.939	Sales, marketing and development managers
121	0.882	Business services and administration managers
231	0.863	University and higher education teachers
413	0.858	Keyboard operators

the index derived above. For Felten et al. (2018) I have used the excel file containing the data for their Appendix A, made available through their github page [6]. Similarly for Eloundou et al. (2023) I have downloaded the file containing the final scores at the occupation level from the associated repository [7]. Finally, for Webb (2022) I've accessed the data through an online notebook made available by the author [8].

As three papers are based on the O*NET occupational classification, I used a crosswalk provided by ESCO to partially match the two sets of occupations. The crosswalk contains a many-to-many mapping of O*NET identifiers to ESCO occupations. The matches are classified as either "broad", "close", "narrow" or "exact". When performing the match, for each of the three papers, the average exposure score was taken for all O*NET occupations corresponding to a given ESCO occupation, regardless of the type of match. Conversely, I have also produced an O*NET version of the exposure score developed in this paper by carrying out the same procedure in reverse - taking an average of all ESCO occupations matching a given O*NET id. This version of the results is also available in the online repository. For Webb (2022) the results were presented as per David Dorn's occupational schema Dorn (2009), so an additional step was needed to match first match them to O*NET IDs. This was done using a crosswalk provided by the author in the online notebook.

After completing the matching, Pearson correlation coefficients were calculated among all four indices first at the individual occupational level and then at the ESCO 3-digit level (as individual occupational scores might be affected by potential noise in the ESCO-to-O*NET matching). The results are presented in Appendix A.4 and Table 3 respectively.

The results of this paper exhibit a strong correlation with Felten's occupational exposure both at the occupation group level - $r = 0.702$, with a Pearson product-moment correlation test yielding $t = 10.531$ on 114 degrees of freedom ($p < 2.2e-16$) - and on the single occupation level - $r = 0.56$, $t = 33.22$, $df = 2359$, $p < 2.2e-16$). Similarly, Eloundou et al's result is also strongly correlated with the index proposed here - $r = 0.66$, $p < 2.2e-16$ on the group level and $r = 0.54$, $p < 2.2e-16$ on the individual occupation level. The relationship is presented graphically in Figure 3 for the group level, and in 8 in the appendix at the more granular individual occupation level.

[6] https://github.com/AIOE-Data/AIOE

[7] https://github.com/openai/GPTs-are-GPTs

[8] https://www.notion.so/michaelwebb/Data-for-The-Impact-of-Artificial-Intelligence-on-the-Labor-Market-3b52b281505a48b8be107d11d8d0c363

Figure 2: AI Capabilities and Research Skills

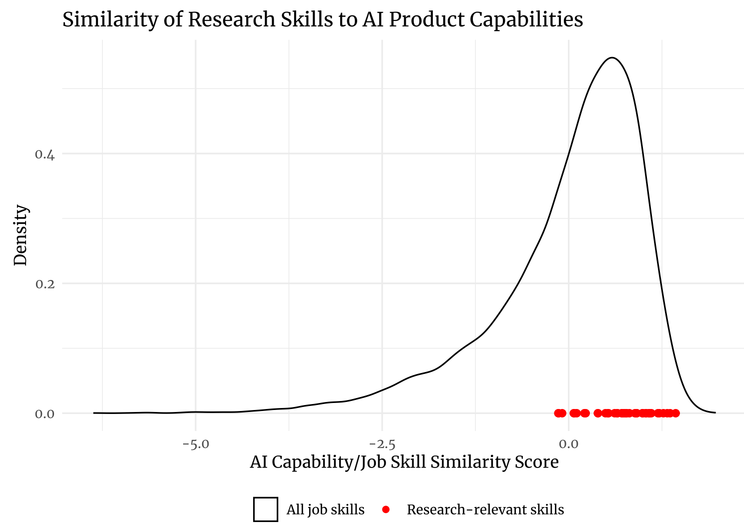
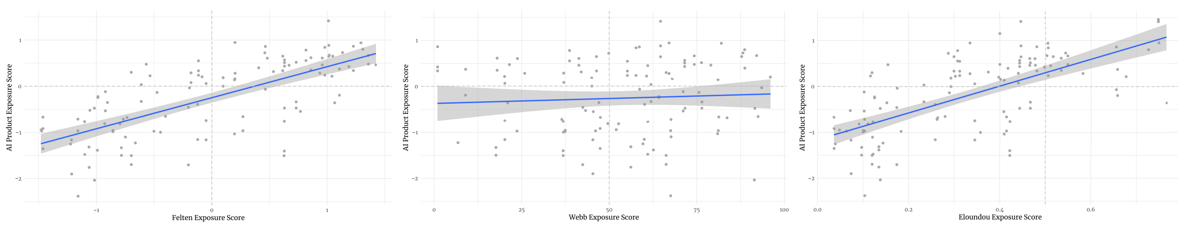


Figure 3: Comparison with Felten et al. (2018), Webb (2022), and Eloundou et al. (2023) at the ESCO 2-digit level



(a) Comparison with Felten et al. (b) Comparison with Webb (c) Comparison with Eloundou

Conversely, the correlation coefficient with Webb's AI score is small and insignificant at the occupation group level ($r = 0.06$, $t = 0.637$, $p = 0.52$). At the individual occupation level, there is significant positive correlation, although it's relatively small in magnitude ($r = 0.106$, $p < 1.5e-6$). Notably, there is no significant correlation between Felten's and Webb's measures, as well as Eloundou's and Webb's.

The indices of Felten et al and Eloundou et al correlate with each other very strongly ($r = 0.86$ at the group level, $r = 0.91$ at the individual level, both significant). The relatively lower correlation that each of them has with the index developed in this work might indicate that the proposed measure successfully captures the difference between what is in principle possible to automate (as measured by Felten et al and Eloundou et al) and what the market actually tries to automate.

To examine the difference with those two measures in more detail, one can look at occupation groups that have a low AI product exposure score and a high Felten/Eloundou score or vice versa. Doing this reveals that both Felten et al and Eloundou et al place relatively high exposure on professions such as secretaries, authors and journalists, and general office clerks. On the other hand, the product-exposure score identifies telecommunication workers, salespeople, and other clerical support workers as relatively more exposed to AI. Example occupations with contrasting product exposure score and Felten/Eloundou scores are presented in Table 7 and Table 8 in the Appendix.

4.4 Relationship to Job Market Outcomes

To better understand the characteristics of occupations with higher exposure to AI product disruption, this section compares the exposure index to various job market outcomes. As discussed above, the list of highly exposed occupations suggests that higher-skilled and higher-paid professions are more likely to be affected by Artificial Intelligence. To confirm this, the exposure score is compared to data on average unemployment and occupational wage premium (available at the ESCO 2-digit level).

The data for the unemployment rate is taken from the CEDEFOP's Skills Intelligence Report European Centre for the Development of Vocational Training (2024). Specifically, the data were gathered using CEDEFOP's online tool for interactively examining the report[9]. The unemployment indicator in the report is calculated as the number of people who became unemployed within the last two years based on their last occupation as a share of the combined number of employed and unemployed in that occupation. The data is available on the EU level and was collected using an automated script directly from the occupation pages of the online report. The resulting graph is shown in the second panel of Figure 4.

The wage premium data is taken from the 2021 Wage Determinants report, which is available through the website of the European Commission European Commission. Statistical Office of the European Union. (2021). The report is based on the 2018 edition of the Structure of Earnings Survey, which collects information about the earning of a representative sample of 11 million workers across the EU. It then builds a set of country-level regression model to explain the variation of earnings. Among the variables used in those regressions are occupation dummies at the 2-digit ISCO level. Since the outcome variable in each regression is the standardized wage per country for that country, the coefficients in each of the country model are on the same scale. Thus, to derive a measure of the relative occupational wage premium at the EU level, I simply take an average of each dummy coefficient across the EU countries. The results are presented in the first panel of Figure 4.

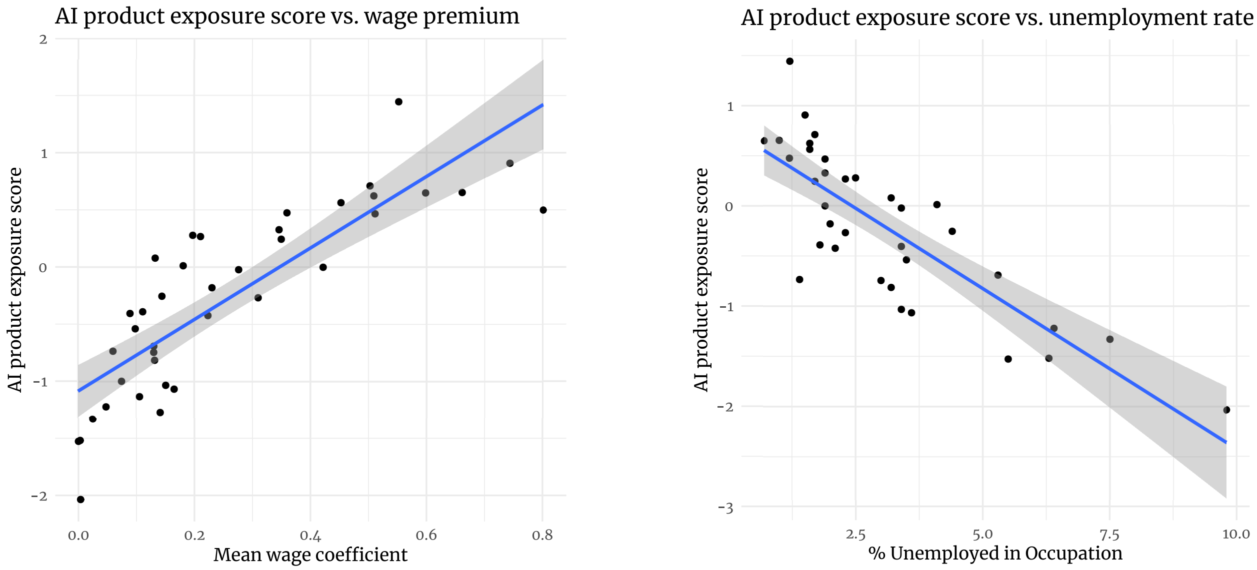
As evident from the plot, the AI product exposure score is strongly negatively correlated with the unemployment rate within the occupation ($r = -0.82$, $t = -8.158$, $df = 33$, $p = 2.04e-9$), and strongly positively correlated with relative wages ($r = 0.84$, $t = 9.38$, $df = 37$, $p = 2.53e-11$). This is perhaps the most intriguing finding of this paper, as previous waves of automation have typically disproportionately affected lower-skilled and lower-paid workers (Acemoglu and Johnson, 2023). One way to interpret these data is that the introduction of AI-powered products reflects not only the underlying expansion of technological capabilities, but also market incentives, as firms try to automate tasks for which labor is scarce and expensive.

It is worth noting that the other indices examined in the previous subsection also exhibit statistically significant positive correlation with wage premiums and negative correlation with unemployment. The unemployment correlation coefficients are -0.61, -0.40, and -0.55 for Felten et al, Webb, and Eloundou et al respectively. The corresponding coefficients for the wage correlation are (in the same order) 0.74, 0.25, and 0.62. They are all significant at the 5% level, with the exception of the correlation between Webb's index and wages ($p = 0.12$). These results are summarized in 9 in the appendix.

The fact that the product exposure score is more strongly correlated with job market outcomes can be interpreted as giving credence to the idea that information about product releases more closely captures actual market implementation of AI relative to other ways to measure exposure.

Another set of data points made available in CEDEFOP's Skills Intelligence tool concern the total number of employees per occupation and the share of women employed in each ISCO level-

[9]Accessed via this URL: https://www.cedefop.europa.eu/en/tools/skills-intelligence



(a) AI Exposure vs Wage Premiums

(b) AI Exposure vs Unemployment

Figure 4: AI Exposure and Job Market Outcomes

2 job. I collect these data in a similar manner to the unemployment figures discussed above and examine their correlation with the AI exposure indices. The results are presented graphically in 9 in the appendix. AI exposure does not seem to correlate with neither the total number of employees nor with the prevalence of women in the occupation. This suggests that AI's impacts will not be skewed to high or low employment jobs, and that we shouldn't expect it to impact women disproportionately.

Before concluding this section, I briefly examine the geographical distribution of occupations highly exposed to AI. Automation-inducing technologies tend to have uneven impacts across regions (Acemoglu and Restrepo, 2020). To identify potential regional imbalances, I calculate a weighted national exposure score using data from CEDEFOP on the number of people employed in each ISCO 2-digit occupation across EU countries. The data are again taken from CEDEFOP's online tool for examining their Skills Intelligence Report. The results are presented in Figure 5.

The countries with an occupational mix most susceptible to AI disruption, as per this measure, are Luxembourg (0.12), Sweden (0.08), and the Netherlands (0.02). In contrast, Romania (-0.36), Bulgaria (-0.33), Spain (-0.27), and Italy (-0.24) have occupational structures that are relatively less exposed to AI [10].

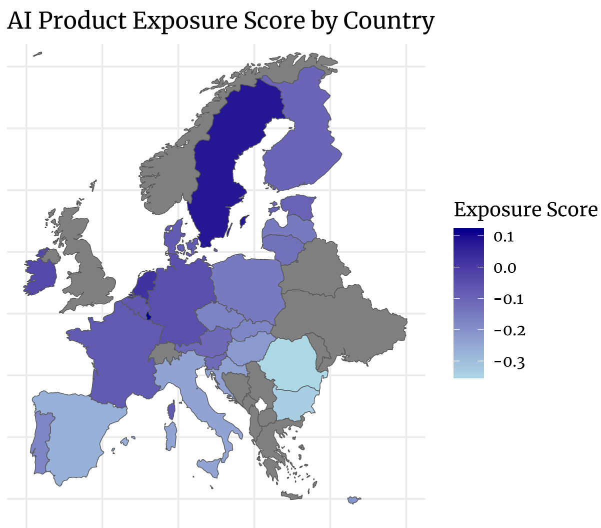
It is worth noting that this is a preliminary look at the exposure of regional labor markets to Artificial Intelligence, considering only the occupational mix and disregarding the demographic structure and overall economic conditions of each region.

5 Conclusion

This study introduces a novel methodology to measuring occupational exposure to artificial intelligence by analyzing corporate press releases about AI product launches and adoptions. This is achieved by employing semantic vector embeddings as well as large language models for text sum-

[10] A more informative analysis would also look at sub-national data; however, I was unable to find regional data on employment by ISCO occupation

Figure 5: Occupational AI Exposure by Country



marization and comparison. By collecting and analyzing this new and rich dataset, the proposed approach is able to identify not only which skills and occupations are in principle susceptible to AI automation, but also which tasks are actively being targeted by market participants.

The results presented in this paper suggest that AI's impact on the labor market may differ significantly from previous technological revolutions, as it is more likely to affect well-paid and prestigious occupations. "Knowledge workers" seem to be most exposed to AI disruption, with occupations such as various types of technology professionals, researchers, educators, health workers, administrators, and managers having the highest exposure index. Furthermore, higher exposure to AI seems robustly correlated to higher wages and lower unemployment.

While largely in agreement with previous studies that aim to develop similar exposure metrics, the index proposed in this paper exhibits stronger correlation with labor market outcomes. This suggests that firms innovate in a directional manner, targeting for automation tasks and occupations that are relatively expensive and in short supply. This lends credence to the idea that using data about product announcements might paint a more accurate picture of the uses and applications of AI compared to other approaches.

As a caveat to the automation interpretation presented above, it is worth underlining that, as with other task-based measures, a higher exposure score for a given occupation does not unambiguously indicate a higher probability of automating said occupation. If we view an occupation as a bundle of tasks, automating some of them might allow allow workers to focus and specialize of the parts of their jobs that are harder for AI to do. In that case, while there could still be potential labor market disruptions, they may take the form of re-bundling of tasks and the creation of new occupations, rather than unemployment.

While it remains to be seen to what extent jobs will be automated and whether new task creation and productivity gains will outpace displacement effects, the occupational exposure score developed in this study provides a valuable tool for future explorations of such questions. The score can be employed as a predictor variable in macro- and micro-economic models of labor displacement, or it may serve to focus efforts of measuring productivity gains or job reorganization.

Two examples of studies that could utilize the exposure measure are provided by recent work of Acemoglu and co-authors. Firstly, Acemoglu et al. (2022) use the metrics developed by Webb and Felten as treatment variables in a macroeconomic model of job automation, tested against data about online job vacancies. They find significant hiring differences in firms with more exposed occupational structures, showing an increase in hiring of professionals likely to implement AI systems and a decrease in hiring for other occupations. However, they do not detect aggregate employment effects.

Similarly, in a related work Acemoglu (2024) employs the index developed by Eloundou et al to derive a big-picture view of the impacts of AI on gross domestic product as a whole. Combining the resulting estimates of the proportion of the economy that could be affected by AI with individual RCT derived measures of productivity improvements due to AI, Acemoglu is able to derive some bounds on the expected total factor productivity growth that can arise due to AI adoption. Using the measure proposed here, which is based on actual product deployments, could provide additional insights and precision to either of these two studies.

Future research could also explore the gap between what is theoretically possible to automate (as measured by e.g. Felten's method) and what is actually implemented (as measured in this study). The strong correlation with wage premiums and unemployment rates indicates that entrepreneurs are most interested in innovating where labor is relatively scarce and expensive. Developing a more robust model of directed technical change and entrepreneurial activity, perhaps within the framework of Aghion and Howitt (1992), could enhance our understanding of this behavior.

Beyond academic research, these findings can inform policymakers as societies adjust to the potential disruptions brought by AI. The occupational exposure scores highlight potential areas of skill mismatch and can help identify targets for retraining programs or other policy interventions. When combined with national and sub-national data on occupational composition, the exposure scores can aid in regional planning and development, as previous impacts of automation technologies have often been concentrated in specific regions or metropolitan areas (Acemoglu and Johnson, 2023).

Furthermore, as highlighted in the Results section, the current wave of AI automation is particularly effective at automating and accelerating parts of research and development. This is significant because research productivity is crucial for long-term growth in endogenous and semi-endogenous growth models, such as those by Romer (1986) and Jones (2022). Investigating how these productivity-enhancing tools affect research output could help us better understand and anticipate AI's impact on the knowledge production function and long-term productivity growth.

In conclusion, while the full impact of AI on the labor market remains to be seen, this study provides a data-driven foundation for understanding which occupations are likely to be at the forefront of AI-driven change. As society navigates this technological transition, continued monitoring and analysis of AI's effects on the labor market will be key for ensuring equitable and productive outcomes.

References

PRNewswire: press release distribution, targeting, monitoring and marketing — prnewswire.com. https://www.prnewswire.com/, 2024. [Accessed 09-06-2024].

D. Acemoglu and S. Johnson. *Power and Progress: Our Thousand-year Struggle Over Technology and Prosperity*. PublicAffairs, 2023. ISBN 9781541702530. URL https://books.google.bg/books?id=ttxfzwEACAAJ.

Daron Acemoglu. The Simple Macroeconomics of AI. 2024.

Daron Acemoglu and David Autor. Chapter 12 - skills, tasks and technologies: Implications for employment and earnings. In David Card and Orley Ashenfelter, editors, *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier, 2011. doi: https://doi.org/10.1016/S0169-7218(11)02410-5. URL https://www.sciencedirect.com/science/article/pii/S0169721811024105.

Daron Acemoglu and Pascual Restrepo. The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, 108(6):1488–1542, June 2018. ISSN 0002-8282. doi: 10.1257/aer.20160696. URL https://pubs.aeaweb.org/doi/10.1257/aer.20160696.

Daron Acemoglu and Pascual Restrepo. Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30, May 2019. ISSN 0895-3309. doi: 10.1257/jep.33.2.3. URL https://pubs.aeaweb.org/doi/10.1257/jep.33.2.3.

Daron Acemoglu and Pascual Restrepo. Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, 128(6), June 2020.

Daron Acemoglu, David Autor, Jonathon Hazell, and Pascual Restrepo. Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1):S293–S340, April 2022. ISSN 0734-306X, 1537-5307. doi: 10.1086/718327. URL https://www.journals.uchicago.edu/doi/10.1086/718327.

Philippe Aghion and Peter Howitt. A model of growth through creative destruction. *Econometrica*, 60(2):323, March 1992. ISSN 00129682. doi: 10.2307/2951599. URL https://www.jstor.org/stable/2951599?origin=crossref.

Philippe Aghion, Benjamin F Jones, and Charles I Jones. Artificial intelligence and economic growth. In Ajay Agrawal, Joshua Gans, and Avi Goldfarb, editors, *The Economics of Artificial Intelligence: An Agenda*, chapter 237-282, page Chapter Page Range. The University of Chicago Press, 2019.

William J. Baumol. Macroeconomics of unbalanced growth: The anatomy of urban crisis. *The American Economic Review*, 57(3):415–426, 1967. URL http://www.jstor.org/stable/1812111.

James Bessen. *Learning by Doing: The Real Connection between Innovation, Wages, and Wealth*. Yale University Press, New Haven, CT, 2015. ISBN 9780300195668.

Erik Brynjolfsson and Tom Mitchell. What can machine learning do? workforce implications. *Science*, 358(6370):1530–1534, December 2017. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.aap8062. URL https://www.science.org/doi/10.1126/science.aap8062.

Erik Brynjolfsson, Tom Mitchell, and Daniel Rock. What can machines learn and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, 108:43–47, May 2018. ISSN 2574-0768, 2574-0776. doi: 10.1257/pandp.20181019. URL https://pubs.aeaweb.org/doi/10.1257/pandp.20181019.

Erik Brynjolfsson, Danielle Li, and Lindsey R. Raymond. Generative ai at work, 2023.

David Dorn. *Essays on Inequality, Spatial Interaction, and the Demand for Skills.* Dissertation, University of St. Gallen, September 2009. Data Appendix, pp. 121-138.

Peter Eckersley and Yomna et al. Nasser. Eff ai progress measurement project. https://eff.org/ai/metrics, 2017-. Accessed on 2017-09-09.

Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. Gpts are gpts: An early look at the labor market impact potential of large language models, August 2023. URL http://arxiv.org/abs/2303.10130. arXiv:2303.10130 [cs, econ, q-fin].

European Centre for the Development of Vocational Training. Skills intelligence tool. Web Page, 2024. URL https://www.cedefop.europa.eu/en/tools/skills-intelligence. Accessed: 2024-06-09.

European Commission, Directorate-General for Employment, Social Affairs and Inclusion. European skills, competences, qualifications and occupations (esco) handbook. European Commission, 2019. URL https://esco.ec.europa.eu/en. Accessed: 2024-06-09.

European Commission. Statistical Office of the European Union. *Wages determinants in the European Union: evidence from structure of earnings survey (SES 2018) data : 2021 edition.* Publications Office, LU, 2021. URL https://data.europa.eu/doi/10.2785/62601.

Edward W. Felten, Manav Raj, and Robert Seamans. A method to link advances in artificial intelligence to occupational abilities. *AEA Papers and Proceedings*, 108:54–57, May 2018. ISSN 2574-0768, 2574-0776. doi: 10.1257/pandp.20181021. URL https://pubs.aeaweb.org/doi/10.1257/pandp.20181021.

Edward W. Felten, Manav Raj, and Robert Seamans. How will language modelers like chatgpt affect occupations and industries? *SSRN Electronic Journal*, 2023. ISSN 1556-5068. doi: 10.2139/ssrn.4375268. URL https://www.ssrn.com/abstract=4375268.

Carl Benedikt Frey and Michael A. Osborne. The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114:254–280, January 2017. ISSN 00401625. doi: 10.1016/j.techfore.2016.08.019. URL https://linkinghub.elsevier.com/retrieve/pii/S0040162516302244.

Matthew Honnibal, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. spacy: Industrial-strength natural language processing in python, 2022. URL https://zenodo.org/record/1212303.

Charles I. Jones. The past and future of economic growth: A semi-endogenous perspective. *Annual Review of Economics*, 14(Volume 14, 2022):125–152, 2022. ISSN 1941-1391. doi: https://doi.org/10.1146/annurev-economics-080521-012458. URL https://www.annualreviews.org/content/journals/10.1146/annurev-economics-080521-012458.

National Center for O*NET Development. O*net online, 2024. URL https://www.onetonline.org/. Retrieved April 24, 2024.

OpenAI. Gpt-4 technical report, 2024.

Paul M. Romer. Increasing returns and long-run growth. *Journal of Political Economy*, 94(5): 1002–1037, October 1986. ISSN 0022-3808, 1537-534X. doi: 10.1086/261420. URL https://www.journals.uchicago.edu/doi/10.1086/261420.

Michael Webb. The impact of artificial intelligence on the labor market, 2022. URL https://www.michaelwebb.co/webb_ai.pdf.

Joseph Zeira. Workers, machines, and economic growth. *The Quarterly Journal of Economics*, 113(4):1091–1117, 1998. URL https://EconPapers.repec.org/RePEc:oup:qjecon:v:113:y:1998:i:4:p:1091-1117.

A Additional Pots and Tables

A.1 Skill-level Exposure Scores

Figure 6: Skill-Level Exposure to AI

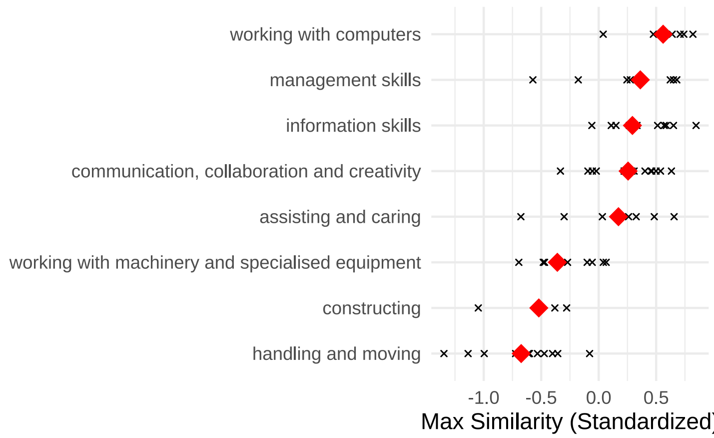


Figure 7: Standardized average of the skill-level maximum similarity with identified AI capabilities. ESCO 1-digit skills on the rows (and as red rhomboids), ESCO 2-digit skills as black crosses.

Table 2: ISCO Skill Groups and Mean Exposure Scores

ISCO 1-digit skill groups	ISCO 2-digit skill groups	Mean Exposure Score
working with computers	programming computer systems	0.741
working with computers	accessing and analysing digital data	0.709
information skills	managing information	0.651
management skills	building and developing teams	0.651
working with computers	setting up and protecting computer systems	0.639
communication, collaboration and creativity	solving problems	0.633
management skills	developing objectives and strategies	0.622
information skills	monitoring developments in area of expertise	0.588
information skills	analysing and evaluating information and data	0.570
communication, collaboration and creativity	liaising and networking	0.538
information skills	processing information	0.512
assisting and caring	counselling	0.483
working with computers	using digital tools for collaboration, content creation and problem solving	0.476

Continued on next page

Table 2 continued from previous page

ISCO 1-digit skill groups	**ISCO 2-digit skill groups**	**Mean Exposure Score**
communication, collaboration and creativity	advising and consulting	0.461
communication, collaboration and creativity	teaching and training	0.454
communication, collaboration and creativity	designing systems and products	0.403
management skills	making decisions	0.373
management skills	leading and motivating	0.358
management skills	organising, planning and scheduling work and activities	0.357
information skills	calculating and estimating	0.333
assisting and caring	protecting and enforcing	0.326
communication, collaboration and creativity	obtaining information verbally	0.308
communication, collaboration and creativity	working with others	0.289
management skills	allocating and controlling resources	0.275
information skills	conducting studies, investigations and examinations	0.266
assisting and caring	providing health care or medical treatments	0.260
communication, collaboration and creativity	presenting information	0.255
management skills	supervising people	0.246
communication, collaboration and creativity	negotiating	0.230
communication, collaboration and creativity	using more than one language	0.221
information skills	documenting and recording information	0.150
information skills	monitoring, inspecting and testing	0.109
working with machinery and specialised equipment	using precision instrumentation and equipment	0.065
working with machinery and specialised equipment	installing, maintaining and repairing electrical, electronic and precision equipment	0.040
working with computers	using digital tools to control machinery	0.039
assisting and caring	providing information and support to the public and clients	0.032
communication, collaboration and creativity	creating artistic, visual or instructive materials	-0.021
communication, collaboration and creativity	writing and composing	-0.055

Continued on next page

Table 2 continued from previous page

ISCO 1-digit skill groups	ISCO 2-digit skill groups	Mean Exposure Score
working with machinery and specialised equipment	operating watercraft	-0.056
information skills	measuring physical properties	-0.060
communication, collaboration and creativity	promoting, selling and purchasing	-0.094
management skills	performing administrative activities	-0.178
working with machinery and specialised equipment	operating aircraft	-0.271
assisting and caring	providing general personal care	-0.301
communication, collaboration and creativity	performing and entertaining	-0.333
handling and moving	handling animals	-0.356
working with machinery and specialised equipment	operating mobile plant	-0.361
working with machinery and specialised equipment	installing, maintaining and repairing mechanical equipment	-0.362
constructing	building and repairing structures	-0.381
handling and moving	making moulds, casts, models and patterns	-0.401
working with machinery and specialised equipment	operating machinery for the extraction and processing of raw materials	-0.471
handling and moving	handling and disposing of waste and hazardous materials	-0.474
working with machinery and specialised equipment	driving vehicles	-0.482
constructing	installing interior or exterior infrastructure	-0.502
handling and moving	tending plants and crops	-0.533
management skills	recruiting and hiring	-0.572
handling and moving	assembling and fabricating products	-0.602
handling and moving	sorting and packaging goods and materials	-0.611
handling and moving	moving and lifting	-0.611
assisting and caring	preparing and serving food and drinks	-0.677
working with machinery and specialised equipment	operating machinery for the manufacture of products	-0.695
handling and moving	transforming and blending materials	-0.723
handling and moving	cleaning	-0.996
constructing	finishing interior or exterior of structures	-1.050

Continued on next page

Table 2 continued from previous page

ISCO 1-digit skill groups	**ISCO 2-digit skill groups**	**Mean Exposure Score**
handling and moving	using hand tools	-1.140
handling and moving	washing and maintaining textiles and clothing	-1.350

A.2 Mean Exposure Scores at the 2-digit Occupation Level

Table 3: Mean AI Exposure Scores - ESCO 3-digit Occupation

ISCO Code	Occupation	Mean Exposure Score
25	Information and communications technology professionals	1.440
12	Administrative and commercial managers	0.907
24	Business and administration professionals	0.709
13	Production and specialised services managers	0.651
22	Health professionals	0.648
23	Teaching professionals	0.622
21	Science and engineering professionals	0.562
11	Chief executives, senior officials and legislators	0.498
32	Health associate professionals	0.474
14	Hospitality, retail and other services managers	0.465
33	Business and administration associate professionals	0.326
41	General and keyboard clerks	0.278
01	Commissioned armed forces officers	0.276
43	Numerical and material recording clerks	0.267
02	Non-commissioned armed forces officers	0.243
31	Science and engineering associate professionals	0.243
03	Armed forces occupations, other ranks	0.183
44	Other clerical support workers	0.078
42	Customer services clerks	0.012
26	Legal, social and cultural professionals	-0.002
34	Legal, social, cultural and related associate professionals	-0.023
74	Electrical and electronic trades workers	-0.180
35	Information and communications technicians	-0.268
53	Personal care workers	-0.254
54	Protective services workers	-0.390
75	Craft and related trades workers	-0.405
72	Metal, machinery and related trades workers	-0.423
52	Sales workers	-0.538
51	Personal service workers	-0.691
61	Market-oriented skilled agricultural workers	-0.734
83	Drivers and mobile plant operators	-0.744
82	Assemblers	-0.813
62	Market-oriented skilled forestry, fishery and hunting workers	-0.997
81	Stationary plant and machine operators	-1.030
71	Building and related trades workers, excluding electricians	-1.070
95	Street and related sales and service workers	-1.130
73	Handicraft and printing workers	-1.270
93	Labourers in mining, construction, manufacturing and transport	-1.220
94	Food preparation assistants	-1.330
96	Refuse workers and other elementary workers	-1.520
91	Cleaners and helpers	-1.530

Continued on next page

Table 3 continued from previous page

ISCO Code	Occupation	Mean Exposure Score
92	Agricultural, forestry and fishery labourers	-2.040

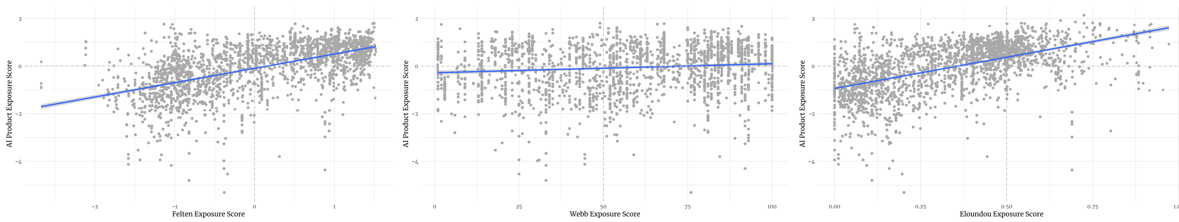
A.3 Lowest Exposure Scores at the 3-digit Occupation Level

Table 4: Lowest Exposure Score Occupations at the 3-digit ISCO Level

ISCO Code	**Exposure Score**	**Occupation**
752	-1.48	Wood treaters, cabinet-makers and related trades workers
512	-1.50	Cooks
731	-1.50	Handicraft workers
951	-1.51	Street and related service workers
813	-1.70	Chemical and photographic products plant and machine operators
962	-1.70	Other elementary workers
814	-1.76	Rubber, plastic and paper products machine operators
911	-1.90	Domestic, hotel and office cleaners and helpers
921	-2.04	Agricultural, forestry and fishery labourers
932	-2.38	Manufacturing labourers

A.4 Correlations With Existing Indices

Figure 8: Occupation-Level Comparison with Felten et al. (2018), Webb (2022), and Eloundou et al. (2023)



(a) Comparison with Felten et al. (b) Comparison with Webb (c) Comparison with Eloundou

Table 5: Correlations with Existing Indices - Individual Occupation Level

	AI Product Exposure Score	Felten et al	Webb	Eloundou et al
Felten et al	0.565 *($p < 0.001$, $n = 2361$)*	–		
Webb	0.106 *($p < 0.001$, $n = 2061$)*	0.032 *($p = 0.146$, $n = 2061$)*	–	
Eloundou et al	0.544 *($p < 0.001$, $n = 2473$)*	0.855 *($p < 0.001$, $n = 2293$)*	-0.0005 *($p = 0.982$, $n = 2061$)*	–

Table 6: Correlations with Existing Indices - ESCO 2-digit Occupation Level

	AI Product Exposure Score	Felten et al	Webb	Eloundou et al
Felten et al	0.702 *($p < 0.001$, $n = 116$)*	–		
Webb	0.060 *($p = 0.525$, $n = 116$)*	-0.058 *($p = 0.535$, $n = 116$)*	–	
Eloundou et al	0.660 *($p < 0.001$, $n = 122$)*	0.908 *($p < 0.001$, $n = 116$)*	-0.053 *($p = 0.571$, $n = 116$)*	–

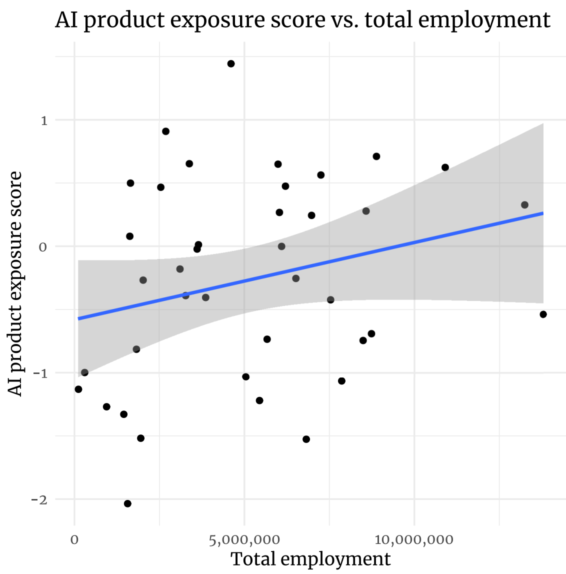
Table 7: Examples of occupations with different AI Product Exposure Score and Felten et al. Exposure Scores. "High" and "Low" are defined as positive and negative scores respectively, as both measures are centered around zero.

High Felten et al Score, Low Product Exposure Score				Low Felten et al Score, High Product Exposure Score			
ISCO	Occupation	Product Score	Felten Score	ISCO	Occupation	Product Score	Felten Score
412	Secretaries (general)	-0.194	1.100	742	Electronics and telecom. installers	0.034	-0.611
264	Authors, journalists and linguists	-0.358	1.060	524	Other sales workers	0.059	-0.050
531	Child care workers and teachers' aides	-0.545	0.766	441	Other clerical support workers	0.078	-0.181
131	Production managers in agriculture, etc.	-0.338	0.729	312	Mining, manufacturing supervisors	0.203	-0.085
621	Forestry and related workers	-0.470	0.729	335	Regulatory government professionals	0.224	-0.535
613	Mixed crop and animal producers	-0.472	0.628	324	Veterinary technicians and assistants	0.237	-0.116
952	Street vendors (excluding food)	-0.758	0.627	432	Material-recording and transport clerks	0.288	-0.181
521	Street and market salespersons	-1.400	0.627	754	Other craft and related workers	0.300	-0.697
951	Street and related service workers	-1.510	0.627	321	Medical and pharmaceutical technicians	0.333	-0.116
411	General office clerks	-0.067	0.463	342	Sports and fitness workers	0.474	-0.568
343	Artistic, cultural and culinary professionals	-0.658	0.268	325	Other health associate professionals	0.548	-0.116

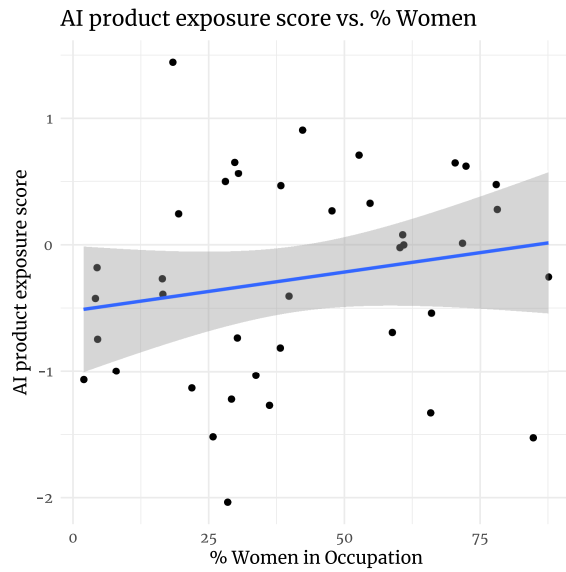
Table 8: Examples of occupations with different AI Product Exposure Score and Eloundou et al. Exposure Scores. For Eloundou et al. "high" and "low" are defined as above and below 0.5 respectively, as their index ranges from 0 (no susceptible tasks) to 1 (all tasks susceptible to automation).

High Eloundou et al Score, Low Product Exposure Score				Low Eloundou et al Score, High Product Exposure Score			
ISCO	Occupation	Product Score	Eloundou Score	ISCO	Occupation	Product Score	Eloundou Score
264	Authors, journalists and linguists	-0.358	0.766	742	Electronics and telecom. installers	0.034	0.234
412	Secretaries (general)	-0.194	0.658	524	Other sales workers	0.059	0.345
411	General office clerks	-0.067	0.655	441	Other clerical support workers	0.078	0.440
				314	Life science technicians	0.137	0.410
				216	Architects, planners, surveyors	0.158	0.402
				031	Armed forces, other ranks	0.183	0.293
				312	Mining, manufacturing supervisors	0.203	0.397
				335	Regulatory government professionals	0.224	0.302
				532	Personal care workers in health	0.230	0.117
				324	Veterinary technicians and assistants	0.237	0.312

A.5 Relationship with Job Market Outcomes



(a) AI Exposure vs Total Employment



(b) AI Exposure vs Percent Women in Occupation

Figure 9: AI Exposure and Additional Job Market Outcomes

Exposure Measure	Wages	Unemployment	Total Employment	Percent Women
AI Product Exposure Score	0.839 *(p<0.001)*	-0.818 *(p<0.001)*	0.256 *(p=0.116)*	0.187 *(p=0.253)*
Felten et al	0.740 *(p<0.001)*	-0.605 *(p<0.001)*	0.155 *(p=0.345)*	0.266 *(p=0.101)*
Webb	0.253 *(p=0.12)*	-0.403 *(p=0.016)*	-0.228 *(p=0.163)*	-0.580 *(p<0.001)*
Eloundou et al	0.617 *(p<0.001)*	-0.552 *(p<0.001)*	0.198 *(p=0.228)*	0.224 *(p=0.171)*

Table 9: Correlation between AI Exposure Scores and Economic Indicators