UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

Dipartimento di Economia
Marco Biagi

DEMB Working Paper Series

N. 272

Becoming The Man Without Qualities? Deskilling in the Age of Artificial Intelligence

Michele Cantarella[1], Giuseppe Molinari[2], Chiara Strozzi[3]

June 2026

[1] Technical University of Denmark - DTU
Email: micant@dtu.dk
[2] Department of Economics Marco Biagi, University of Modena and Reggio Emilia
Email: giuseppe.molinari@unimore.it
[3] Department of Economics Marco Biagi, University of Modena and Reggio Emilia, and RECent, IZA, LISER
Email: chiara.strozzi@unimore.it

ISSN: 2281-440X online

Becoming The Man Without Qualities? Deskilling in the Age of Artificial Intelligence

MICHELE CANTARELLA[1], GIUSEPPE MOLINARI[*2], and CHIARA STROZZI[2,3]

[1]Technical University of Denmark - DTU
[2]University of Modena and Reggio Emilia
[3]IZA@LISER

Abstract This paper investigates how Artificial Intelligence reshapes the human capabilities that jobs require. Using longitudinal O*NET data for the U.S. labour market over 2011–2025, we distinguish among three types of human capabilities - abilities, skills, and knowledge - and construct two measures of human capabilities' exposure to AI: one based on observed progress in Generative AI benchmark performance and one based on the broader evolution of AI-related scientific and public attention. We document a dual pattern. Within occupations, greater AI exposure is associated with higher proficiency requirements for selected capabilities. At the occupational level, more exposed occupations exhibit a compression in the overall breadth of capabilities required. Together, these findings suggest that AI is driving a process of occupational restructuring, leading to more specialized and less diverse capability profiles embedded in jobs.

Keywords: Artificial Intelligence; AI exposure; Skill reallocation; Task content; Deskilling.

JEL Codes: J24; J21; O33.

---

[*] Corresponding author. Email: giuseppe.molinari@unimore.it

1 Introduction

The rapid progress of Generative AI has renewed debates about the future of work. Much of the debate has focused on a central question: whether AI will destroy jobs or create new ones. Pessimists point to the displacement potential of AI, while optimists emphasise its complementarity with human labour, invoking the historical resilience of labour demand in the face of technological change. Yet this debate obscures a deeper transformation. As AI systems take over an increasing range of tasks, workers may gradually lose the opportunity to exercise and maintain the full breadth of their capabilities, a process more akin to deskilling than to displacement, even within occupations that survive the advent of AI.

The possibility that technological change may erode workers' capabilities has long been central to debates on deskilling and the evolution of work under technological and organisational change. A foundational perspective on this debate is provided by Braverman (1974), who interprets deskilling as an organisational process driven by managerial control. Building on Scientific Management principles (Taylor, 1911), this view argues that the fragmentation of production processes and the separation of conception from execution progressively reduce workers' autonomy and dilute embodied knowledge (Marglin, 1974; Thompson, 1983).

While Braverman emphasised organisational control and the labour process, subsequent economic research approached similar concerns through the lens of technological change, skill demand, and task composition. The literature on skill-biased technological change (SBTC) argues that technological progress shifts labor demand toward more educated workers (Katz and Murphy, 1992; Acemoglu, 2002), helping explain the rise in education and skill wage premiums observed since the 1980s (Autor et al., 1998; Goldin and Katz, 2007). Moving from macro-skills to specific activities, the routine-biased technological change (RBTC) framework posits that technology substitutes for routine tasks while complementing abstract cognitive work (Autor et al., 2003; Acemoglu and Autor, 2011). This task-based perspective has been used to explain employment polarisation, as digital technologies automate routine tasks while complementing analytical and interpersonal activities (Goos and Manning, 2007; Goos et al., 2009, 2014). Later refinements further examine how technology modifies task bundles along dimensions such as autonomy, social interaction, and codifiability (Spitz-Oener, 2006; Fernández-Macías and Hurley, 2017). However, subsequent evidence complicates this linear interpretation: Beaudry et al. (2016) document that, following the tech bust, high-skilled workers increasingly moved down the occupational ladder, displacing less educated workers in lower-skill jobs. This reversal in cognitive task demand suggests that technological change may compress the occupational structure from the top, rather than merely hollowing out the middle, thereby blurring the boundaries between upskilling and deskilling.

While these two traditions focus either on organizational control or on the evolution of task bundles, our approach shifts the focus to the underlying human capabilities required to perform those tasks. Rather than asking how Artificial Intelligence alters external task structures or changes overall labor demand, we examine how AI advances are associated with changes in the abilities, skills and knowledge. By analyzing the micro-level prerequisites of work rather than its macro task composition, we offer a distinct perspective on the upskilling-deskilling debate, capturing how technology reshapes what workers must actually know and be able to do on the job. Building on this perspective, we ask whether exposure to AI is associated with changes in the importance-weighted proficiency levels of human capabilities within occupations, capturing both the importance of a capability and the level of proficiency required to perform a job. We examine this question by tracing how much these proficiency levels have changed between 2011 and 2025, and then measuring how advancements in AI are associated with these changes.

To address this question, we use data from O*NET,[1] which provides detailed information on the evolution of job-related human capabilities across a large set of occupations in the U.S. labour market. Human capabilities are mapped along three domains: abilities, skills, and knowledge.[2] *Abilities* are enduring attributes of individuals that influence performance; *skills* are developed capacities that facilitate learning or task execution; and *knowledge* refers to organised sets of principles and facts applying to general domains. We combine O*NET data with two distinct datasets to map capability exposure to AI. These two datasets capture: (i) improvements in Generative Artificial Intelligence model performance in specific domains, measured through standardized benchmarks; and (ii) the salience of these advances, which we measure using data from news and research aggregators. The first measure captures realised progress in generative AI performance, while the second provides a broader historical account of how AI-related scientific and public attention has evolved. Both measures are constructed at the capability level and then mapped to the corresponding human capability. We produce results at both the capability and occupational levels. At the former level, we find a positive association: within occupations, the capabilities most exposed to AI experience an increase in the required proficiency. However, once capabilities are aggregated within occupations, AI exposure is associated with a decline in the overall breadth of proficiency require- ments, pointing to a compression of the human content of work. This divergence suggests that AI operates through a dual mechanism: while it enhances those abilities, skills, and knowledge that are more directly complementary to the technology, fostering within-occupation specialisation, it is simultaneously associated with a contraction of the broader capabilities embedded in occupations.

Our paper contributes to the literature in three ways. First, we contribute to the broader literature on AI and the labour market by reconnecting it to the question of deskilling. While much of the existing literature focuses on the employment and task- level consequences of technological change (Acemoglu and Zilibotti, 2001; Autor et al., 2006, 2008; Acemoglu and

---

[1] National Center for O*NET Development. The O*NET content model. O*NET Resource Center. Retrieved May 14, 2026, available at: `https://www.onetcenter.org/content.html`.

[2] National Center for O*NET Development. O*NET OnLine Help: Scales, ratings, and standardized scores. O*NET OnLine. Retrieved May 20, 2026, available at: `https://www.onetonline.org/help/online/scales`.

Autor, 2011; Autor, 2013), we examine how AI is associated with changes in the human capabilities that occupations require. Rather than studying the evolution of tasks within jobs, we focus on the evolution of abilities, skills, and knowledge, thereby providing a capability-based perspective on the upskilling–deskilling debate.

Second, we contribute to the emerging literature measuring AI exposure through human capabilities and occupational requirements (Felten et al., 2019, 2021; Tolan et al., 2021; OECD, 2025, 2026). Unlike existing approaches, we do not impose ex ante classifications of tasks, occupations, or capabilities. Instead, we adopt a data-driven approach in which occupations are characterised by detailed profiles of abilities, skills, and knowledge derived from O*NET. This allows us to identify which capabilities are more exposed to AI and how their relative importance evolves over time. By shifting the unit of analysis from occupations to capabilities, we show that AI exposure may cut across conventional occupational categories and affect specific abilities, skills, and knowledge domains within otherwise different jobs. Existing evidence generally finds that AI exposure is concentrated in highly educated, white-collar occupations, while lower-skill manual jobs appear comparatively less affected (Georgieff and Hyee, 2021; OECD, 2023). Other studies emphasise heterogeneous effects across occupations, ranging from displacement in high-exposure roles to complementarities that enhance the productivity of highly skilled workers (Bessen, 2018; Milanez, 2023; Engberg et al., 2025). This perspective suggests that AI exposure is not solely an occupational characteristic but also a capability-specific phenomenon that cuts across conventional occupational boundaries.

Third, we contribute to the growing literature measuring labour-market exposure to Generative AI (Eisfeldt et al., 2023; Eloundou et al., 2023; Auer et al., 2024; Humlum and Vestergaard, 2024; Massenkoff and McCrory, 2026). While existing approaches typically infer exposure from expert assessments, task classifications, occupational characteristics, firm-level workforce composition, or observed AI adoption, we construct a capability-level measure based on the observed evolution of AI performance across a large set of benchmarks. By tracking improvements in AI capabilities over time, our measure captures technological progress directly rather than relying on static assessments of what AI can do at a given point in time.

The rest of the paper is organised as follows. Section 2 describes the data. Section 3 explains the methodology. Section 4 outlines the empirical strategy and presents the first set of results. Section 5 discusses the main findings. Section 6 concludes. The Appendix provides additional details on specific steps of the analysis, the construction of the dataset, and further descriptive evidence.

2 Data

Our empirical analysis draws on three distinct types of information: data on the development of Artificial Intelligence, data on the evolution of proficiency levels across occupations, and labour market data. Our unit of analysis is an occupation–human capability cell in the first stage, and an occupational-level cell in the second stage.[3] Occupations are classified according to the O*NET-SOC 2019 taxonomy, using the six-digit level of disaggregation. Human capabilities are drawn from the O*NET Content Model[4] and comprise abilities, skills, and knowledge. Each occupation is then matched with occupation-level data from the U.S. Bureau of Labor Statistics' Current Population Survey (CPS), in order to weight each occupation according to its share in the U.S. labour market. Our analysis covers the period from 2011 to 2025.

2.1 Data on AI development

To capture information about AI development, we rely on two datasets: *LLM Stats* and *AI Topics*. *LLM Stats*[5] is a dataset that tracks recent advances in Artificial Intelligence by compiling a large collection of standardized evaluations across multiple AI modalities. It enables systematic comparisons across the main large language models currently available on the market—such as models developed by OpenAI, Anthropic, Google, and other leading providers—by assessing their performance on a wide range of benchmark tasks. Benchmarks consist of standardized evaluation tasks designed to measure model performance in specific domains, such as reasoning, mathematics, language understanding, coding, or multimodal processing. By aggregating performance across these benchmarks, the dataset provides a structured way to compare models along multiple dimensions and to track progress in different areas of Artificial Intelligence (see Appendix for details) over time by mapping each benchmark with the date of release of a model.

AI Topics[6] is a long-standing repository curated by the Association for the Advancement of Artificial Intelligence (AAAI) [7] which compiles a vast range of material related to Artificial Intelligence research, applications, and contributors. The platform aggregates a wide spectrum of AI-related content, including news items, blog posts, conference proceedings, journal articles, and other documents spanning the period from 1905 to 2026, collected automatically through the NewsFinder system (Buchanan et al., 2013). This dataset has been employed in previous empirical analyses of technological change, such as Tolan et al. (2021). The resulting dataset effectively tracks, over time, how salient advances in AI are by capturing the evolution of related academic discourse and media coverage.

---

[3] For example, in the first stage we observe cells such as *Economist–Active Learning* or *Economist–Critical Thinking* for each year from 2011 to 2025. In the second stage, all abilities, skills, and knowledge items are then recomposed and aggregated at the occupational level (e.g. *Chief Executives*).

[4] The Content Model was developed using research on job and organizational analysis. It embodies a framework that captures both the characteristics of occupations (through job-oriented descriptors) and those of workers (through worker-oriented descriptors). In addition, it allows occupational information to be applied across jobs, sectors, and industries (cross-occupational descriptors), as well as within occupations (occupation-specific descriptors). Source: `https://www.onetcenter.org/content.html`, last accessed April 23, 2026.

[5] See `https://llm-stats.com/`.

[6] See `https://aitopics.org/`.

[7] See `http://www.aaai.org/`.

2.2 Human capabilities data

To extract information on human capabilities, we rely on the O*NET Content Model (National Center for O*NET Development, b). The model provides a structured framework to describe both the characteristics of occupations and those of workers, integrating job-oriented descriptors with worker-oriented attributes. In addition, it allows for the analysis of information both across occupations—through cross-occupational descriptors—and within occupations, through more detailed occupation-specific elements. These descriptors are organised into a set of three domains, i.e., *abilities, skills,* and *knowledge*, that capture the key dimensions of work. O*NET data are collected using a multi-method approach based on surveys administered to a statistically representative sample of workers and establishments in the U.S. labour market. Firms are first sampled, and workers within those firms are then randomly selected to provide information on job requirements and work activities. The database is regularly updated to reflect changes in occupational content, making it particularly suitable for analysing the evolution of human capabilities over time. For each human capability, O*NET provides two complementary measures: an *importance* score and a *level* score. The importance score captures the relevance of a given descriptor for a specific occupation, while the level score measures the degree of proficiency that a descriptor requires to perform the job (National Center for O*NET Development, a).[8] Together, these measures allow us to quantify both the relative significance and the depth of each human capability within occupations. By weighting the level by the importance, we construct our proficiency level variable.

2.3 Labour market data

Data used to describe the U.S. labor market are drawn from the Current Population Survey (CPS), sponsored jointly by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics (BLS). The CPS is the primary source of labor force statistics for the United States and provides information for the period 2011–2025 considered in our analysis. We use CPS employment shares to weight each occupation according to its relative size in the U.S. labor market. In addition, we extract a set of individual and occupational covariates used as controls, including industry, education, age, gender, and wage levels.

3 Measuring human capabilities' exposure to AI

In this section, we describe how we construct a new measure of the exposure of human capabilities to Artificial Intelligence using two distinct datasets, showing how we link this measure to the observed variation in abilities, skills, and knowledge.

---

[8] For instance, according to the latest O*NET update, the occupation *Economist* scores 4.4 on the 0–7 level scale and 4.1 on the 1–5 importance scale.

Most of the existing literature measures exposure to AI through a task-based approach, mapping AI benchmarks or capabilities to job tasks (Brynjolfsson et al., 2018; Gmyrek et al., 2023; Eloundou et al., 2024), or by relying on patent data to infer technological exposure (Webb, 2019). While these approaches capture how technology reshapes the content of work, they remain anchored to the objective dimension of jobs, i.e. the task composition of occupations. In contrast, we adopt a capability-oriented perspective that focuses directly on the evolution of human capabilities, i.e. abilities, skills, and knowledge, thereby capturing how the composition of capabilities, rather than only jobs, adjusts to technological change. In this respect, our approach is related to Felten et al. (2021) and Tolan et al. (2021), who move beyond task-based mappings by linking AI capabilities to cognitive abilities. However, our framework departs from these contributions in an important way. Existing studies typically focus on a restricted subset of abilities that are ex ante considered relevant for AI exposure, which may introduce a degree of endogeneity in the selection of worker characteristics. By contrast, we construct a comprehensive mapping that includes the full set of human capabilities defined in O*NET, linking each ability, skill, and knowledge item to AI developments.

This approach allows us to capture the entire capability profile of workers across occupations, including those dimensions that may be less intensively used in a given job but still contribute to the overall structure of human capital. As a result, our measure reflects not only how AI interacts with core job requirements, but also how it relates to the broader composition and evolution of worker capabilities in the labour market.

3.1 AI exposure

In this section, we describe the construction of our AI exposure measures. The first measure is based on LLM Stats and captures improvements in the performance of major generative AI models. The second measure is based on AI topics and captures the salience of advances in Artificial Intelligence more broadly.

3.1.1 LLM Stats–based AI exposure

The AI exposure measure based on LLM Stats is particularly informative because it captures a distinct dimension of Artificial Intelligence. Unlike measures derived from broader indicators of AI research activity, this index is constructed directly from the performance of Large Language Models across a range of benchmark categories, including reasoning, general capabilities, vision, multimodal tasks, mathematics, agentic behaviour, language, and coding. A more detailed description of the benchmarks is provided in the Appendix A2.

From LLM Stats we download all 383 available AI benchmarks, each of which reports the performance of major models on a task-specific evaluation metric. For each benchmark and each company, we select the highest score achieved among all models released by that company. This identifies the technological frontier marked by the release model (e.g., OpenAI, Anthropic, and others). Since data are available only for the period 2023–2025, we aggregate, for each year, the average of these frontier values across all companies to obtain an annual benchmark-level performance measure.

Because all benchmark scores are already expressed on a 0-1 scale, the resulting dataset assigns to each benchmark a directly comparable value of machine-learning capability for each available year. The next step is to link each benchmark to a specific human capabilities. To link each benchmark to a specific human capabilities, we compare the textual descriptions of benchmarks with the O*NET definitions of Abilities, Skills, and Knowledge. We construct a semantic similarity measure linking textual descriptions of AI capabilities (benchmarks) to human capabilities (abilities, skills, and knowledge described in the O*NET database). The goal is to obtain a continuous score capturing how closely each AI benchmark corresponds to a specific human capabilities. Conceptually, the approach relates to recent text-based matching methods connecting technological content to occupational characteristics (Montobbio et al., 2022, 2024).

Let $B = \{b_1, \ldots, b_{N_B}\}$ denote the set of benchmarks and $O = \{o_1, \ldots, o_{N_O}\}$ the set of O*NET elements. For each benchmark b_i, we concatenate its name and description into a single text string $T(b_i)$, and analogously for each O*NET element we use the combined descriptive text $T(o_j)$. All texts undergo minimal normalization consisting of lowercasing and removal of irregular whitespace, with no stemming, lemmatization, or stopword filtering, since such operations tend to distort semantic representations.

Each text is then mapped into a dense semantic vector using a pre-trained Sentence-BERT model (Reimers and Gurevych, 2019), specifically the *all-MiniLM-L6-v2* architecture widely used in semantic textual similarity applications (Yin and Zhang, 2024). Denoting the embedding of a text x as $E(x)$, we L2-normalize all vectors to obtain $\hat{E}(x) = E(x)/\|E(x)\|$. Semantic similarity between benchmark b_i and O*NET element o_j is computed as cosine similarity between their normalized embeddings:

$$S_{ij} = \hat{E}(b_i) \cdot \hat{E}(o_j) = \sum_{k=1}^{d} \hat{E}(b_i)_k \, \hat{E}(o_j)_k,$$

where d is the embedding dimension. Since vectors are normalized, $S_{ij} \in [-1, 1]$ and measures the conceptual proximity of the two textual descriptions. Higher values of S_{ij} indicate that the linguistic and semantic content of the benchmark closely resembles that of the corresponding human capability item. The resulting mapping dataset consists of tuples (b_i, o_j, S_{ij}) for all j in the selected top-K set for each i. These scores provide a continuous measure of the conceptual alignment between AI benchmarks and human capabilities. Importantly, the similarity score S_{ij} is used only to decide whether benchmark b_i is meaningfully related to O*NET element o_j. It is not used to determine the size of the AI exposure assigned to that element. After constructing the benchmark–human capability links, we assign AI exposure to each O*NET element by transferring the exposure value of the matched benchmarks to the corresponding human capabilities. When multiple benchmarks are associated with the same capability, we aggregate benchmark-level exposure by taking their average, obtaining a single exposure value for each O*NET element. Thus, semantic similarity defines the mapping between AI benchmarks and human capabilities, while benchmark performance determines the magnitude of exposure.

As in the case of the AI Topics index, we then translate these human capabilities values into an occupation-level measure. For each occupation, the AI exposure of every human capability is weighted by its O*NET Importance value, so that requirements that are more central to an occupation contribute proportionally more to its overall exposure. Finally, the resulting occupation–human capabilities scores are standardised on a 0–1 scale to ensure comparability across occupations and across years.

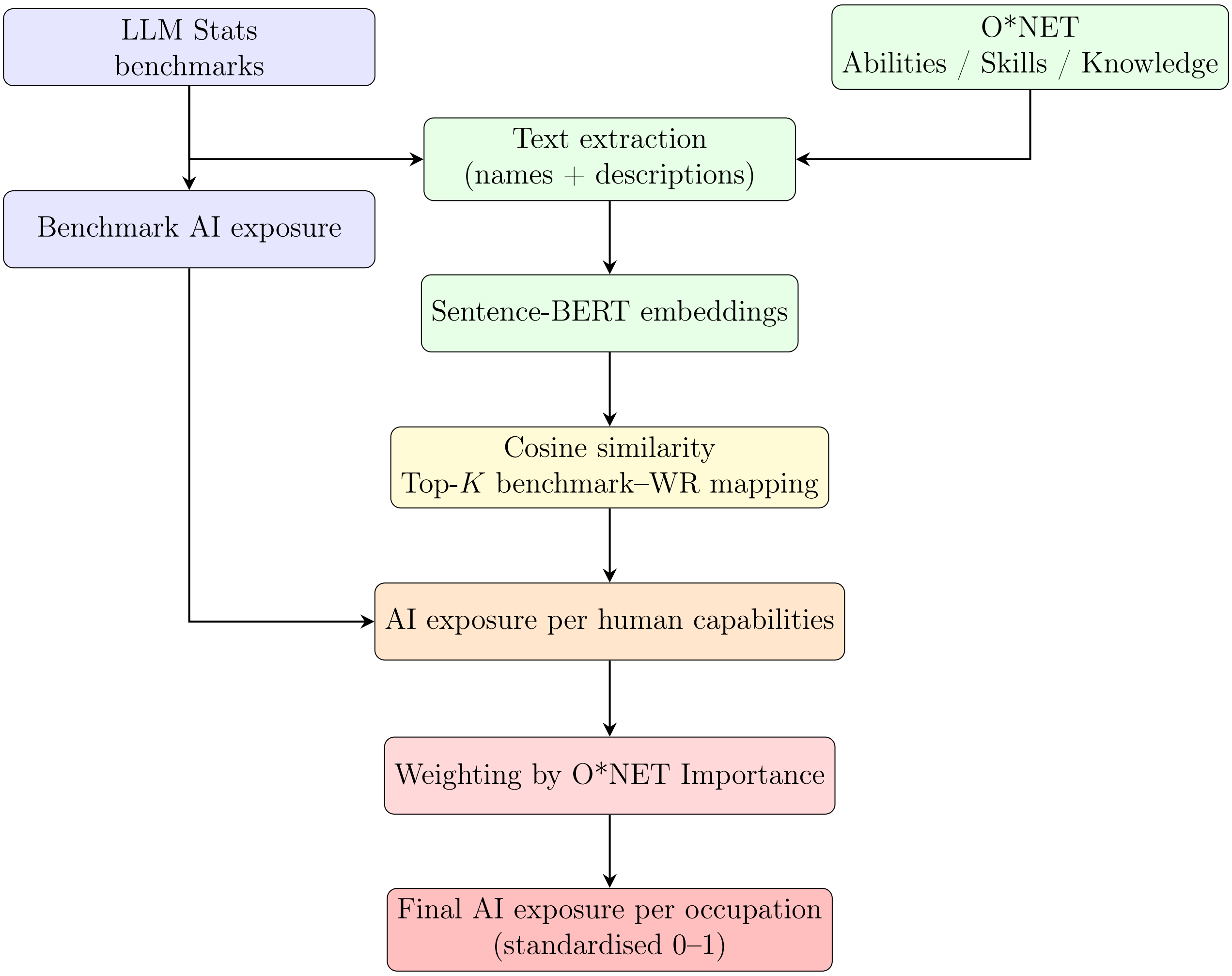

Figure 1: **Workflow for building the LLM Stats–based AI exposure**

3.1.2 AI Topics–based AI exposure

This subsection describes the construction of the AI exposure measure based on the *AI Topics* dataset. The objective is to quantify the extent to which developments in Artificial Intelligence relate to specific human capabilities. While the benchmark-based measure captures realised improvements in model performance, the *AI Topics* measure captures the research and media salience of AI developments associated with each capability.

The construction proceeds as follows. For each human capability, we extract its textual definition from O*NET and use it to build a query that captures the relevant conceptual domain. Each query is manually expanded to include synonyms and semantically related expressions to retrieve a comprehensive set of AI-related publications.

For example, for the Ability *Inductive Reasoning*, defined as "the ability to combine pieces of information to form general rules or conclusions," the corresponding query is:

```
"inductive reasoning" OR ("combine" AND "information") OR ("form" AND
"conclusions") OR "pattern recognition".
```

Using these queries, we collect, for each year between 2011 and 2025, the number of AI-related publications associated with each human capability. To capture the accumulation of AI knowledge over time, we construct a cumulative publication measure, defined as the total number of AI-related publications associated with a given capability up to year t. This approach reflects the idea that technological advances build on previous research and that the stock of knowledge available at a given point in time is more informative than annual publication flows alone.

The resulting cumulative publication counts are then standardised to a 0–1 scale using a global transformation. This procedure accounts for both temporal variation in AI research activity and structural differences across capabilities, as some domains naturally attract more research output than others. The outcome is an AI index defined at the level of human capabilities, where each ability, skill, or knowledge item is assigned a unique AI exposure value.

This capability-level AI exposure is then weighted by the corresponding O*NET Importance score within each occupation. As a result, the contribution of a given capability to the overall AI exposure of an occupation depends on its relative importance in that occupation. For instance, occupations in which *Inductive Reasoning* is less relevant will receive a lower contribution from this capability compared to occupations where it plays a central role.

Table 1: **Procedure for the construction of the AI exposure measure from AI Topics**

Step	**Description**
1. Identify worker requirements	Select all Abilities, Skills, and Knowledge items from the O*NET Content Model.
2. Extract O*NET definitions	Retrieve the textual definition of each worker requirement to build a semantic search query.
3. Construct AI Topics query	For each worker requirement, construct a Boolean query including key terms, synonyms, and conceptual expansions, e.g. *Inductive Reasoning.*
4. Retrieve yearly publication counts	Query AI Topics for each year (2011–2025) and extract the number of AI-related publications matching the query.
5. Global standardisation	Standardise publication counts to a 0–1 scale across all worker requirements and years to capture temporal trends and absolute differences in research volume.
6. Construct worker-requirement AI index	Assign each worker requirement a yearly AI score based on its standardised publication count.
7. Weight by occupational importance	Multiply the worker-requirement AI score by O*NET *Importance* values for each occupation, generating an occupation–worker-requirement–year AI exposure measure.

Notes: The procedure is applied to all worker requirements in the O*NET Content Model and repeated annually. The resulting index captures the intensity of AI-related research associated with each worker requirement over time.

3.2 Human capabilities

Human capabilities play a dual role in our empirical strategy. First, as discussed in the section on AI exposure, they are used to construct our measures of exposure to AI. Second, they define the dependent variable of the analysis, namely the change in each human capability between 2011 and 2025. To construct our dataset, we begin by assembling, for each occupation–human capability pair, the corresponding *importance*[9] and *level*[10] scores provided directly by O*NET.

As a result, the occupation *Economist* features roughly 120 human capabilities observations per year, multiplied by 15 years, yielding about 1,800 observations. Accounting for missing values, the final dataset comprises 1,645,906 observations organised in occupation–human capabilities–year cells, with each cell containing an Importance and a Level score, both ranging from 0 to 100.

We then construct a key variable in which the *Level* score is weighted by its *Importance* score. This approach is widely used in the literature since Autor et al. (2003). We focus on the Level score because it captures a qualitative dimension of human capability: it measures *how*

---

[9] *Importance* measures how critical a given ability, skill, or knowledge area is for performing a specific occupation.

[10] *Level* measures the intensity, complexity, or degree of proficiency required for a given ability, skill, or knowledge area in a specific occupation.

much of a certain attribute an individual needs to perform a task, and therefore how Artificial Intelligence may affect that requirement. By weighting Level by Importance, we incorporate the extent to which each human capability matters for the occupation under consideration.

Given that O*NET is updated multiple times per year[11], we average the Importance and Level scores within each calendar year to obtain a single yearly value. This choice also mitigates inconsistencies in the database, since some updates produce erratic or highly noisy scores. Finally, because O*NET presents important limitations for panel construction, we do not exploit the full yearly variation as a balanced occupation–capability panel. Instead, we use the yearly averages to construct our final outcome as the long-difference in the importance-weighted level of each human capability between 2011 and 2025. Formally, for each occupation o and human capability h, we define:

$$HC_{oht} = Level_{oht} \times Importance_{oht}$$

and construct the dependent variable as:

$$\Delta HC_{oh}^{2011-2025} = HC_{oh,2025} - HC_{oh,2011}$$

where HC_{oht} denotes the importance-weighted level of capability h in occupation o at time t. This measure captures whether the required proficiency of a given human capability within an occupation has increased or decreased over the period considered.

Because O*NET does not update all occupations every year, some occupation-capability pairs are missing observations in either 2011 or 2025. In these cases, we apply an iterative procedure: if the 2025 value is unavailable, we use the closest subsequent observation (2024, 2023, and so on); similarly, if the 2011 value is unavailable, we use the closest available observation after 2011 (2012, 2013, etc.). This approach allows us to recover meaningful changes while minimizing the loss of observations.

To avoid attenuation bias in the estimated capability changes, we exclude occupations whose most recent update of either level or importance occurred before 2013. For these occupations, capability requirements remain mechanically constant over most of the sample period, generating changes close to zero that may reflect outdated information rather than genuine stability in occupational requirements. We also remove observations with a level equal to zero in either the initial or final year, as these requirements are effectively absent from the occupation and therefore unlikely to exhibit meaningful variation over time.

Finally, we rely on simple differences rather than growth rates. Percentage changes would inflate small variations for low-level human capabilities while underweighting equally meaningful changes for those starting from medium or high levels.

---

[11] See `https://www.onetcenter.org/dataUpdates.html` for a complete list of updates.

3.3 Data integration and harmonization

Before proceeding with the analysis, we merge the dataset structured at the *occupation–human capabilities–year* level with labor market information from the *Current Population Survey* (CPS). The CPS is used for two purposes. First, it provides employment shares, which we use to weight occupations when aggregating results. Second, it provides occupation-level control variables included in Equation (2), such as wages, industry composition, educational attainment, and the demographic composition of workers. Since CPS microdata are coded using the 2010 six-digit Standard Occupational Classification (SOC) system, whereas O*NET relies on the O*NET-SOC taxonomy, we harmonize the two sources using the SOC 2010–O*NET-SOC crosswalk provided by Hardy et al. (2018). This step is necessary because the two datasets rely on different occupational classification systems and the same occupational title may not correspond one-to-one across them. The crosswalk allows us to assign CPS-based labour market variables, including employment shares and occupation-level controls, to the corresponding O*NET occupations. When multiple O*NET-SOC occupations correspond to the same SOC code, CPS variables are assigned to all matched O*NET occupations. In this way, the O*NET information on human capabilities can be consistently combined with CPS information on labour market structure and worker characteristics.

4 Empirical strategy and descriptive evidence

4.1 Empirical strategy - capability-occupation level estimates

We now explore the relationship between occupational exposure to Artificial Intelligence and the evolution of Human Capabilities between 2011 and 2025. To do so, we estimate the following regression at the capability–occupation level:

$$y_{co} = \beta X_{co} + \gamma_o + \varepsilon_{co}, \tag{1}$$

where c indexes human capabilities (Abilities, Skills, and Knowledge) and o indexes occupations. Each observation, therefore, corresponds to a specific capability–occupation pair, such as *Active Learning–Economist* or *Active Listening–Physician.* The dependent variable y_{co} is defined as the change between 2011 and 2025 in the standardized, importance-weighted required proficiency level of capability c within occupation o. Thus, the analysis relies on a long-difference specification that captures the overall evolution of capability requirements over the period. The variable x_{co} denotes the AI exposure measure associated with capability c, weighted by its occupational importance in occupation o, using either the LLM-stats or the AI Topics measure. The term γ_o denotes occupation fixed effects, and standard errors are clustered at the capability level. The change in each human capability is computed as the difference between its standardized, importance-weighted level in 2025 and 2011.

The variable X_{co} measures the potential exposure of each human capabilities-occupation unit (c, o) to AI, as described in Section 3. As already discussed, these measures capture the extent to which a given human capability is exposed to Artificial Intelligence within an occupation. We rely on two alternative measures for AI-exposure X_{co}, as described in Section 3, one based on salience, and the other based on the capabilities of AI itself.

A few considerations are in order. The first one concerns the endogeneity of the AI exposure shock. While it is reasonable to expect that the AI shock could not be anticipated, it could be argued that the shock is endogenous to the demand for specific capabilities. To address these issues, in Appendix 7 we provide robustness results where AI exposure is instrumented with a shift-share design defined at the capability level. The instrument combines the pre-AI importance structure of occupations, predetermined with respect to subsequent AI progress, with the AI exposure within each occupation. Reassuringly, the estimates remain consistent with our baseline.

A second consideration concerns the importance weighting. Both the dependent and the explanatory variable are weighted by O*NET importance, which could in principle induce a mechanical association between them. Two features of the construction limit this concern. First, the two variables weight importance for different objects: the dependent variable is a long difference in the importance-weighted level of a capability, while the regressor is an importance-weighted measure of its exposure to AI. A common importance factor scaling a change on one side and a level on the other does not generate a mechanical correlation of fixed sign, since the change can be positive or negative independently of the importance weight. Second, we introduce occupation fixed effects to absorb any occupation-level component of importance, so that identification rests on within-occupation, across-capability variation rather than on differences in average importance across occupations.

The last consideration concerns the interpretation of the estimated coefficient. Since our focus is on how human capabilities evolve over time, these indices are not designed to capture the degree of complementarity or substitutability between AI and tasks or occupations, in the spirit of Pizzinelli et al. (2023). Rather, they are intended to measure a more "intangible" dimension of exposure.

The underlying idea is that the more exposed a human capability is to AI, the more likely it is to be reshaped, regardless of whether the occupation as a whole is substituted or complemented by AI. In other words, an occupation may continue to grow in employment while its internal capability profile changes, with some capabilities being required at higher levels of proficiency and others becoming less central.

We therefore interpret the coefficient β in equation (1) as capturing whether AI-exposed capabilities experience larger changes in their required proficiency within occupations. A positive value of β indicates that capabilities with higher exposure to AI experience larger increases in their standardized, importance-weighted required proficiency between 2011 and 2025. Conversely, a negative value of β implies that more AI-exposed capabilities tend to experience a decline in their required proficiency over time.

4.2 Empirical strategy - occupation level estimates

In addition to the capability–occupation level specification, we also estimate a second model in which all variables are aggregated at the occupation level. Formally, we estimate:

$$y_o = \beta X_o + \Gamma' Z_o + \varepsilon_o, \tag{2}$$

where o indexes occupations. The dependent variable y_o measures the change between 2011 and 2025 in the aggregate level of human capabilities required within occupation o, computed by combining the standardized and importance-weighted capability measures described above. The variable X_o denotes the occupation-level AI exposure index, obtained by aggregating capability-specific exposure measures within each occupation using occupational importance weights. The vector Z_o includes a set of occupation-level controls measured at the beginning of the period, including the initial level of worker requirements, educational composition, age structure, demographic characteristics, and sector fixed effects. Standard errors are heteroskedasticity-robust.

In this specification, β captures the relationship between AI exposure and changes in the overall breadth of human capabilities required within occupations. A positive value of β indicates that more AI-exposed occupations experience an expansion in their aggregate capability requirements, while a negative value of β indicates a narrowing or compression of the capability profile required within those occupations. This specification, therefore, provides a complementary perspective: rather than focusing on how specific capabilities evolve within occupations, it examines whether more AI-exposed occupations become, on average, broader or narrower in terms of the breadth of human capabilities they require over time.

4.3 Descriptive Evidence

This section provides descriptive statistics for the AI exposure measures and documents the evolution of Human Capabilities over time. Table 2 reports summary statistics for the two measures of Artificial Intelligence exposure (the variable x), at the human capabilities-by-occupation level.

Table 2: **Summary Statistics for AI scores (Human Capabilities–Occupation Level)**

Variable	**Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
AI LLM	78,117	0.230	0.120	0	1
AI Topics	74,147	0.214	0.090	0	0.82

Notes: The table reports summary statistics for the two AI score at the human capabilties–occupation level.

The two measures are available for 78,177 and 74,147 cells in our dataset. The AI LLM index displays a slightly higher average value than the alternative measure, consistent with the fact that the latter has a lower upper bound, with a maximum of 0.82.

Table 3 reports the same set of summary statistics, but in this case, computed at the occupation level.

Table 3: **Summary Statistics for AI scores (Occupation Level)**

Variable	**Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
AI LLM	797	0.350	0.180	0	1
AI Topics	797	0.390	0.150	0	1

Notes: The table reports summary statistics for the two AI scores at the occupation level.

The descriptive statistics show clear differences between the AI exposure measures computed at the human capabilities–occupation level and those obtained after aggregating the dataset at the occupation level.

The first set of values (78,117 and 74,147 observations) corresponds to the exposure indices calculated for each human capability within each occupation. These measures reflect how strongly individual Abilities, Skills, and Knowledge elements are exposed to AI. At the capability–occupation level, the average exposure is 0.23 for the AI exposure measure constructed from LLM benchmark performance and 0.21 for the AI exposure measure constructed from AI Topics.

After aggregating the dataset to the occupation level, we compute an occupation-level AI exposure by averaging the human capabilities exposure values within each occupation. The distribution of the indices changes as a mechanical result of this aggregation and because the two measures rely on different numerical scales: the mean AI_LLM exposure at the occupation level is 0.35, while the AI Topics measure shows a substantially higher mean of 0.39.

Table 4 reports the correlations between the two AI exposures at each level of aggregation and shows that a human capabilities–occupation level the two measures display a relatively low correlation, reflecting the conceptual difference between the LLM-based (narrow AI) and the AI-Topics (general AI) indicators. Once we aggregate the data to the occupation level, however, the correlation between the two measures increases substantially. This suggests that part of the difference between the indices dissipates when moving from fine-grained human capabilities to broader occupational aggregates.

Table 4: **Correlation between AI scores**

	Correlation (AI LLM, AI Topics)
Human capabilities–occupation level	0.243
Occupation level	0.786

Notes: The table reports the correlation between the two AI scores computed at different levels of aggregation. The substantially higher correlation at the occupation level reflects the smoothing effect of aggregation across human capabilities.

Tables 5 and 6 report the five most and least exposed human capabilities for both AI exposures. For the LLM-based measure, the most exposed human capability is *Deductive Reasoning*, followed by other communication- and comprehension-related abilities such as *Oral Comprehension*, *Speech Recognition*, and *Reading Comprehension.* These are all highly cognitive abilities typically associated with high-skill occupations, consistent with recent evidence on the types of tasks most affected by AI (Lassébie and Quintini, 2022; Felten et al., 2021). Conversely, the least exposed capabilities correspond to physical or domain-specific knowledge areas, such as *Dynamic Strength*, *Dynamic Flexibility*, and *Food Production*, which are less directly related to the types of tasks currently targeted by large language models.

By contrast, the AI Topics measure identifies *Stamina* as the most exposed human capability, followed by attributes such as *Speech Clarity*, *Originality*, and *Control Precision.* According to O*NET, stamina refers to "the ability to exert yourself physically over long periods of time without getting winded or out of breath," a feature more commonly associated with exposure to broader technological families such as automation or robotics. This pattern suggests that the AI Topics index captures not only cognitive forms of AI exposure, but also the growing research and media attention devoted to the automation of physical processes. In this sense, the topic-based measure points to a broader technological frontier, where AI developments are increasingly connected to robotics, physical tasks, and embodied forms of automation. By contrast, the LLM-based index remains more closely associated with cognitive, language, and reasoning-intensive capabilities.

Table 5: **Human Capabilities with the Highest and Lowest AI LLM Score**

Human Capability	**AI LLM**
Panel A: Highest AI LLM Exposure	
Deductive Reasoning	0.691
Oral Comprehension	0.430
Speech Recognition	0.367
Active Listening	0.341
Reading Comprehension	0.330
Panel B: Lowest AI LLM Exposure	
Dynamic Flexibility	0.001
Dynamic Strength	0.001
Sociology and Anthropology	0.001
Food Production	0.002
Chemistry	0.002

Notes: The table reports the five human capabilities with the highest and lowest exposure according to the LLM Stats-based measure.

Table 6: **Human Capabilities with the Highest and Lowest AI Topics Score**

Human Capability	**AI Topics**
Panel A: Highest AI Topics Exposure	
Stamina	0.600
Speech Clarity	0.385
Originality	0.384
Management of Material Resources	0.381
Control Precision	0.350
Panel B: Lowest AI Topics Exposure	
Sociology and Anthropology	0.001
Negotiation	0.007
Repairing	0.007
Psychology	0.007
Fine Arts	0.008

Notes: The table reports the five human capabilities with the highest and lowest exposure according to the AI Topics-based measure.

We also construct the corresponding tables after aggregating the data at the occupation level. The resulting rankings, reported in Tables 8 and 9, show the most and least exposed occupations according to the two AI exposure measures once human capabilities are collapsed to the occupational dimension.

Table 7: **Highest and Lowest Occupation-Level Exposure According to AI LLM**

Panel A: Highest AI LLM Exposure

Occupation	**AI LLM**
Mathematicians	1.000
Statisticians	0.890
Mathematical Science Teachers, Postsecondary	0.865
Biostatisticians	0.858
Actuaries	0.842

Panel B: Lowest AI LLM Exposure

Occupation	**AI LLM**
Dishwashers	0.017
Fence Erectors	0.016
Roof Bolters, Mining	0.007
Foundry Mold and Coremakers	0.006
Pressers, Textile, Garment, and Related Materials	0.000

Notes: The table reports the five occupations with the highest and lowest values of AI LLM, the occupation-level AI exposure index obtained by aggregating worker requirement exposures.

Table 8: **Highest and Lowest Occupation-Level Exposure According to AI Topics**

Panel A: Highest AI Topics Exposure

Occupation	**AI Topics**
Mathematicians	1.000
Financial Quantitative Analysts	0.922
Statisticians	0.874
Astronomers	0.870
Mathematical Science Teachers, Postsecondary	0.868

Panel B: Lowest AI Topics Exposure

Occupation	**AI Topics**
Dishwashers	0.071
Meat, Poultry, and Fish Cutters and Trimmers	0.069
Cleaners of Vehicles and Equipment	0.025
Fast Food and Counter Workers	0.016
Pressers, Textile, Garment, and Related Materials	0.000

Notes: The table reports the five occupations with the highest and lowest values of the AI Topics index. Higher values indicate greater exposure based on topic-model similarity to AI-related domains.

Although the distribution of workers requirements is centered close to zero, Figure 2 highlights a substantial dispersion in the evolution of human capabilities between 2011 and 2025. The presence of pronounced tails indicates that several skills experienced meaningful increases or decreases over time. This variability confirms that, despite an overall stable core, occupations underwent significant internal adjustments in their capability composition.

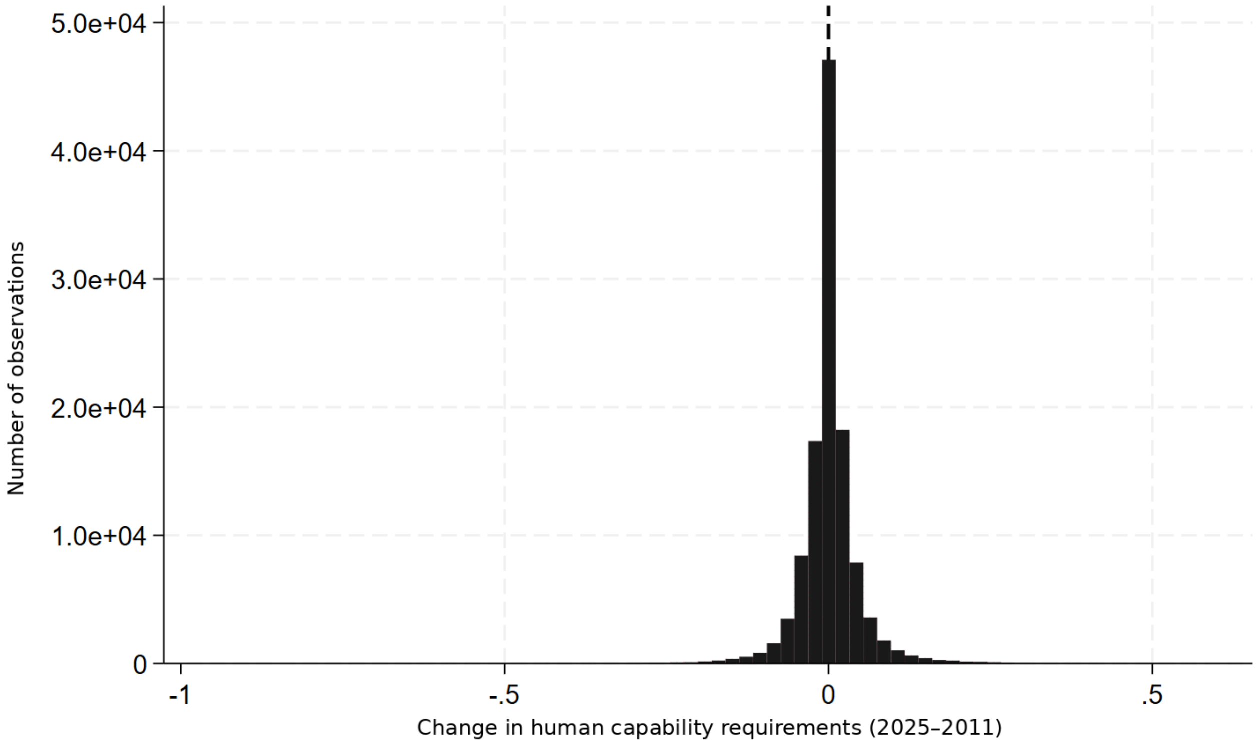

Figure 2: **Distribution of Changes in Human Capabilities , 2011–2025**

Note: Positive values indicate increases in required human capabilities levels, while negative values indicate decreases.

To better capture how human capabilities have evolved across occupations, we aggregate SOC major groups into four macro-occupational categories (cognitive/high-skilled; social and service; sales and office; manual and physical). The pronounced blue-to-red gradient across human capabilities confirms that both the direction and magnitude of changes differ substantially between occupational domains. Across all groups, the heatmap reveals a clear predominance of red shading, indicating widespread declines in required skills, abilities, and knowledge between 2011 and 2025.

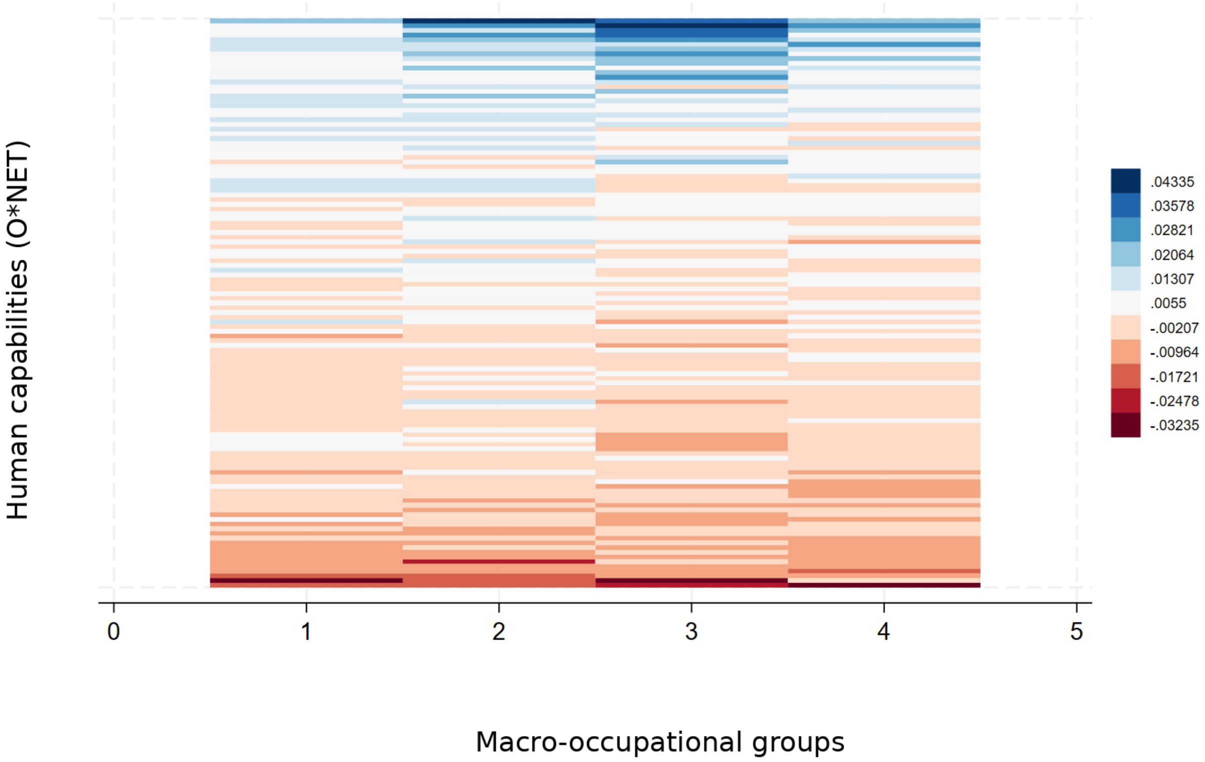

Figure 3: **Heatmap of Changes in human capabilities by Macro-Occupational Group (2011–2025)**

Note Macro-occupational groups are constructed by aggregating SOC major groups as follows: (1) cognitive/high skilled (code 11, 13, 15, 17, 19, 23, 25, 27); (2) social and service (code 21, 29, 31, 33, 39); (3) sales and office (code 41, 43); (4) manual and physical occupations (code 35, 37, 45, 47, 49, 51, 53). Each row represents an ONET ability, knowledge, or skill, ordered from the largest to the smallest average change in human capabilities between 2011 and 2025. Colors indicate the mean change in human capabilities for each group combination

5 Results

5.1 Results by occupation–human capabilities level

Table 9 reports the estimates from regression (1). Baseline estimates are offered in column (1), while column (2) adds occupation fixed effects for robustness. We find a positive and statistically significant association between changes in human capabilities over time and the exposure of human capabilities to Artificial Intelligence within occupations. This result is consistent across both the exposure index constructed from LLM statistics and the broader AI measure derived from AI Topics.

In economic terms, the magnitude of the effect is non-negligible. According to the LLM-based AI exposure indicator, a 0.25-point increase in the exposure index is associated with a 0.5-percentage-point increase in the importance-weighted proficiency required for a given human capability between 2011 and 2025. Using the AI Topics measure, the corresponding increase is approximately 0.44 percentage points, indicating a slightly weaker but still economically meaningful effect when considering a broader notion of AI exposure.

These results suggest that AI exposure is associated with changes in the required proficiency of human capabilities within occupations. Rather than reducing the role of human capabilities, higher AI exposure is associated with certain abilities, skills, and knowledge being required at higher levels of proficiency. This points to a more concentrated capability profile within occupations, which is consistent with a process of within-occupation reconfiguration.

This interpretation is consistent with recent evidence showing that Generative AI can enhance worker performance, facilitate learning, and disproportionately benefit less experienced workers, suggesting that AI often operates as a complement to human capabilities rather than a pure substitute (Brynjolfsson et al., 2025).

Whereas previous studies have documented an increase in tasks compatible with AI's current capabilities (Acemoglu et al., 2020), our results suggest that this transformation extends beyond tasks to the underlying human capabilities themselves. This interpretation is consistent with evidence from McElheran et al. (2023), which shows that AI adoption is concentrated among more educated and innovation-oriented individuals. Taken together, these findings suggest that AI does not uniformly affect all capabilities, but instead selectively enhances those that are more complementary to its development.

Table 10 provides further confirmation of these regression results when we split the sample into Abilities, Skills, and Knowledge. In all three categories, the estimated coefficients remain positive, indicating that the relationship between AI exposure and the evolution of human capabilities is robust across different types of occupational attributes.

However, some heterogeneity emerges in the magnitude of the effects. Using the LLM-based measure, the estimated coefficient remains broadly similar for Abilities, while it becomes larger for Skills and Knowledge. By contrast, when using the AI Topics measure, the estimated coefficients tend to be smaller across all three categories. The lower coefficients observed in the subsamples compared to the full sample likely reflect composition effects, as the aggregate estimate captures cross-category variation that is no longer present when the sample is split.

Table 9: **2011–2025 Change in Human Capabilities**

	(1)	**(2)**
AI LLM	0.020***	0.019***
	(0.0055)	(0.0052)
Observations	78,117	78,117
AI Topics	0.017***	0.016***
	(0.001)	(0.0018)
Observations	74,147	74,147

Source: Authors' calculations.
Notes: The dependent variable measures the change in worker-requirement levels (Abilities, Skills, and Knowledge) between 2011 and 2025. Standard errors are clustered at the worker-requirement level and reported in parentheses. Regression (1) includes occupation fixed effects (O*NET-SOC 6-digit), while regression (2) excludes occupation fixed effects.
Significance levels: $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

Table 10: **2011–2025 Change in Human Capabilities by Capability Domain**

	Abilities (A)	**Knowledge (K)**	**Skills (S)**
AI LLM	0.019***	0.061**	0.037***
	(0.003)	(0.019)	(0.008)
Observations	32,674	21,928	25,728
AI Topics	0.017***	0.04***	0.032***
	(0.001)	(0.010)	(0.002)
Observations	32,674	18,066	23,407

Source: Authors' calculations.
Notes: The dependent variables measure changes in worker-requirement levels between 2011 and 2025, separately for Abilities, Skills, and Knowledge. Standard errors are clustered at the worker-requirement level and reported in parentheses.
Significance levels: $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

In Figures 4 and 5, we further disaggregate the analysis by estimating separate regressions for the most common human capabilities and display the results using forest plots. In these figures, the red dots represent the estimated coefficients, while the horizontal whiskers denote the confidence intervals. The plots confirm that the positive impact of Artificial Intelligence (according to both exposure indices) holds for many individual human capabilities.

The only negative or near-zero effects emerge for *physical abilities*—which is unsurprising, as firms adopting AI rely more heavily on digitized information and cloud computing, suggesting complementarities with other "enabling" technologies (McElheran et al., 2023; Kapoor and Teece, 2021)—and, more interestingly, for *system skills* (which include, for example, Judgment and Decision Making) and for some *process skills* (such as Critical Thinking, Active Learning, and Learning Strategies). This pattern appears consistent with the perspective advanced by Gerlich (2025), although this interpretation is not fully supported in Figure 4, which reports the estimates based on the AI–LLM exposure index.

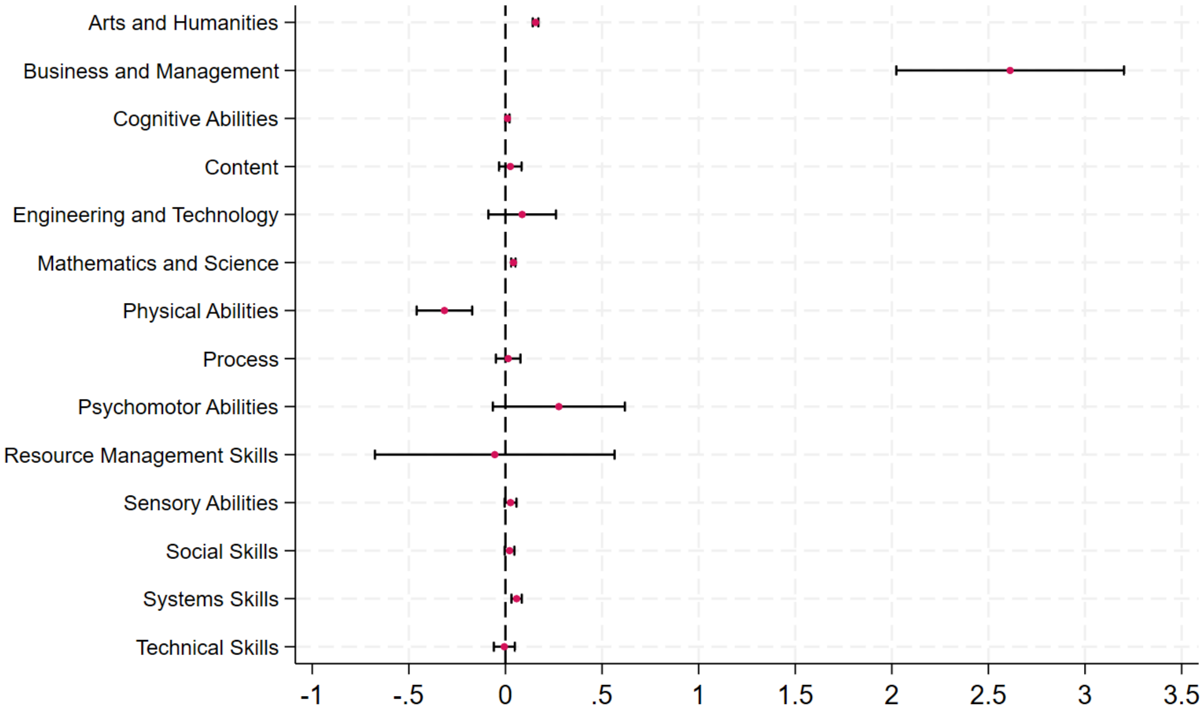

Figure 4: **Forest plot of LLM-Based AI exposure effects on individual human capabilities**

Note: Each red dot represents the estimated coefficient from a separate regression of the change in a given human capability on the AI–LLM exposure index, controlling for occupation fixed effects and clustering standard errors at the human-capability level. Horizontal whiskers show the 95% confidence intervals. Only human capabilities with a sufficient number of observations are included. Positive estimates indicate stronger increases in required human-capability levels between 2011 and 2025 for human capabilities more exposed to the LLM-based AI measure.

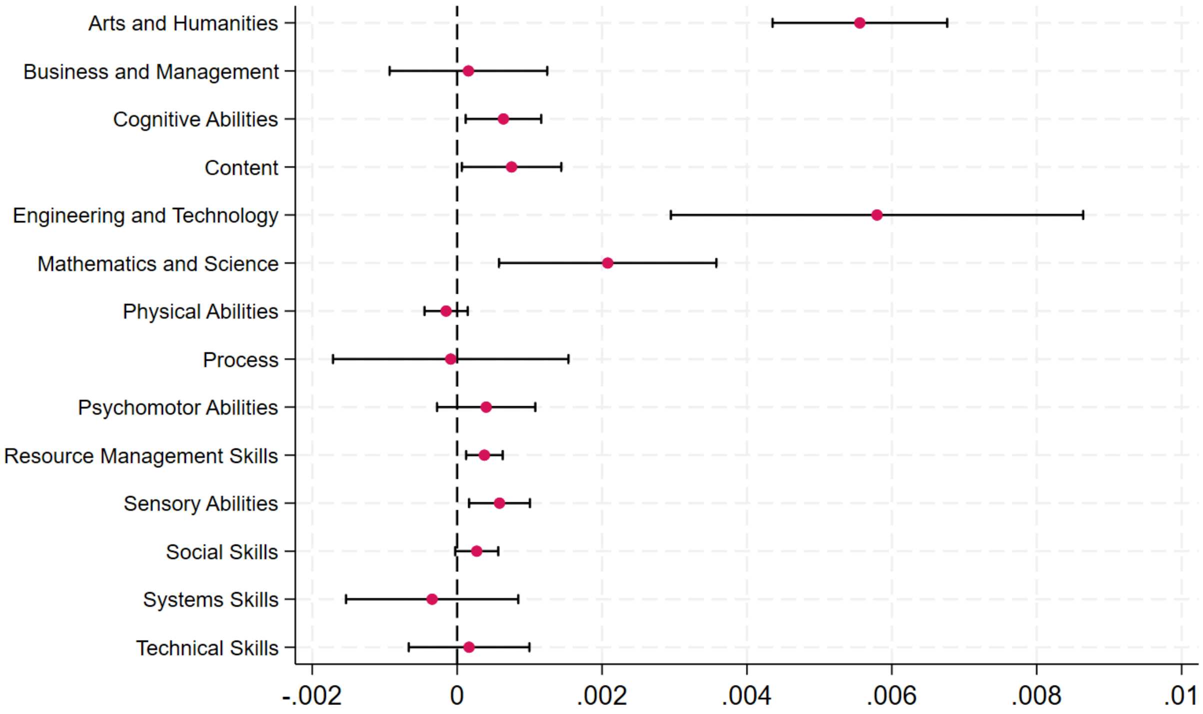

Figure 5: **Forest plot of AI–Topics exposure effects on individual human capabilities**

Note: Each red dot represents the estimated coefficient from a separate regression of the change in a given human capabilities on the AI–Topics exposure index, controlling for occupation fixed effects and clustering standard errors at the human capabilities level. Horizontal whiskers show the 95% confidence intervals. Only human capabilities with a sufficient number of observations are included. Positive estimates indicate stronger increases in required human capabilities levels between 2011 and 2025 for human capabilities more exposed to AI Topics.

5.2 Results at the occupational level

We aggregate the data at the occupational level to examine whether occupations that are more exposed to AI also experience an expansion in the breadth of human capabilities they require. As shown in Table 11, which reports in the first column the estimates based on the full set of human capabilities and in the subsequent columns the results for Abilities, Knowledge, and Skills separately, the picture changes substantially. In this case, the coefficients become negative.

In economic terms, an increase of 0.25 in the AI exposure index is associated with a decrease in human capabilities of about 6.5 percentage points when using the AI Topics indicator and about 3.5 percentage points when using the LLM-based exposure index.

This result stands in sharp contrast with the positive effects previously documented at the occupation–human capability level. At the capability level, AI exposure is associated with increases in the required proficiency of specific abilities, skills, and knowledge. By contrast, once human capabilities are aggregated at the occupational level—thereby capturing the overall breadth of capabilities required within each occupation—the relationship becomes negative.

Taken together, these findings point to a dual dynamic. On the one hand, AI exposure is associated with higher required proficiency in a subset of human capabilities. On the other hand, these increases are accompanied by a reduction in the overall breadth of capabilities required within occupations. In other words, AI appears to reinforce a narrower set of human capabilities while reducing the diversity of capabilities embedded in occupations.

This dynamic may have important implications for labour-market adjustment, as recent evidence suggests that workers exposed to AI differ substantially in their ability to adapt to occupational transitions following displacement (Manning and Aguirre, 2026). More broadly, these findings go beyond the standard interpretation of AI as either a substitute for or a complement to human labour.

Our findings can be interpreted in light of recent work linking task-level technological change to aggregate outcomes. Acemoglu (2025) argues that the macroeconomic effects of AI may remain limited when productivity gains are concentrated at the task level and do not fully translate into aggregate improvements. At the same time, the task-based framework of Autor and Thompson (2025) shows how automation can reshape occupations by altering their expertise requirements and the composition of human labour.

Our results connect these perspectives. At the capability–occupation level, AI exposure is associated with increases in the required proficiency of specific human capabilities, consistent with task-level adjustments. Yet, once these changes are aggregated at the occupational level, the relationship becomes negative, indicating a compression in the overall breadth of human capabilities required within occupations.

This suggests that AI-driven task reallocation may simultaneously enhance certain dimensions of expertise while reducing the overall diversity of capabilities required within occupations. In turn, this mechanism may help explain why significant micro-level improvements do not necessarily translate into large macroeconomic gains.

Table 11: **2011–2025 Change in Human Capabilities at the Occupational Level**

	All Capabilities	**Abilities (A)**	**Knowledge (K)**	**Skills (S)**
AI LLM	-0.140**	-0.120**	-0.018	-0.077
	(0.065)	(0.057)	(0.058)	(0.143)
Observations	372	372	372	372
AI Topics	-0.260***	-0.130**	-0.090	-0.039
	(0.065)	(0.059)	(0.059)	(0.102)
Observations	372	372	372	372

Source: Authors' calculations.

Notes: The table reports cross-sectional estimates of the change in human capabilities between 2011 and 2025 at the occupational level. All specifications control for baseline human capability levels, occupational composition by education, age, gender, and race, and include controls for fifteen macro-industry shares. Robust standard errors are reported in parentheses.

Definitions: A = Abilities; K = Knowledge; S = Skills.

Significance levels: $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

5.3 Robustness checks

A potential concern is that the estimated relationship may be sensitive to the specific construction of the AI exposure measures. Both exposure indicators rely on a number of modelling choices that may affect the intensity assigned to each occupational capability. To assess the robustness of our findings, we reconstruct both measures using alternative aggregation procedures that capture different dimensions of technological progress. If the main results were driven by a particular measurement choice, the estimated coefficients would be expected to change substantially across specifications. Conversely, similar estimates would suggest that the relationship is not an artefact of the baseline construction and reflects a more general association between AI exposure and changes in occupational capabilities. For the AI LLM exposure measure, we construct an alternative indicator that gives greater weight to frontier technological performance. In the baseline specification, annual AI performance is measured by averaging scores across all benchmarks available in a given year. As a robustness check, we instead select, for each year, the best-performing benchmark score and use it to construct the exposure measure. The remaining steps follow the procedure described in the methodology section: benchmark-level AI performance is linked to human capabilities through semantic similarity and then aggregated at the occupation-capability level.

For the AI Topics exposure measure, we similarly modify the temporal construction of the indicator. While the baseline measure is based on the cumulative stock of AI-related research associated with each human capability, the alternative specification is constructed using the growth in AI-related research over the period 2011–2025. This version therefore captures the pace of expansion in AI research activity rather than its accumulated volume. The results are reported in Table 12. Overall, the estimates provide reassuring evidence that the main findings are not driven by the specific construction of the AI exposure measures. At the occupation-capability level, both alternative measures remain positive and statistically significant. The coefficient for AI LLM is close to the baseline estimate, suggesting that the association between LLM-related technological progress and changes in human capabilities is not sensitive to whether AI performance is measured through average benchmark scores or frontier benchmark performance. The coefficient for AI Topics is larger than in the baseline specification, mainly reflecting the different scaling of the growth-based measure, which captures the pace of expansion in AI research rather than its accumulated stock.

The robustness check also supports the occupation-level results. When the alternative exposure measures are aggregated at the occupation level, both AI LLM and AI Topics remain negative and statistically significant.

As an additional robustness check, we examine whether the results depend on the inclusion of occupational importance in the construction of the dependent variable. The baseline analysis uses the Importance–Level measure, which combines the proficiency level required for a capability with its importance within the occupation. This choice reflects the idea that changes in capabilities that are central to an occupation should receive greater weight than changes in capabilities that play only a marginal role. To assess the sensitivity of the results, we replicate the analysis using changes in capability levels alone, excluding the importance component from the dependent variable.

The results are reported in Table 13. The estimated coefficient for AI LLM remains positive and statistically significant at conventional levels close to the 10% threshold, although it is less precisely estimated than in the baseline specification. By contrast, the AI Topics measure becomes statistically insignificant. Overall, these findings suggest that the relationship between AI exposure and capabilities change is stronger when capability requirements are measured using the combined Importance–Level index. This is consistent with the interpretation that AI primarily affects the capabilities that are most relevant within occupations, rather than generating uniform changes in proficiency levels across all capabilities.

Table 12: **Alternative AI Exposure Measures**

Occupation-Capability Level	
AI LLM	0.016***
	(0.0051)
Observations	78,117
AI Topics	0.068***
	(0.0170)
Observations	74,147
Occupation Level	
AI LLM	-0.120***
	(0.057)
Observations	372
AI Topics	-0.067***
	(0.016)
Observations	372

Source: Authors' calculations.

Notes: The dependent variable measures the change in human-capability requirements between 2011 and 2025. At the occupation-capability level, the dependent variable is the change in capability requirements within occupations. At the occupation level, the dependent variable is the change in the aggregate human-capability index. AI LLM is constructed using the highest-performing benchmark observed in each year rather than the average benchmark score. AI Topics is constructed using the growth rate of AI-related publications associated with each human capability rather than cumulative publication counts. Standard errors are reported in parentheses.

Significance levels: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 13: **2011–2025 Change in Human Capability Levels**

	Coefficient
AI LLM	0.008*
	(0.0045)
Observations	78,117
AI Topics	0.002
	(0.0042)
Observations	74,147

Source: Authors' calculations.
Notes: The dependent variable measures the change in capability levels between the baseline year and the final year. Unlike the baseline specification, which uses the combined Importance–Level measure, this robustness check considers changes in capability levels only. Standard errors are clustered at the worker-requirement level and reported in parentheses.
Significance levels: $*p < 0.10$; $**p < 0.05$; $***p < 0.01$.

6 Conclusion

This paper examines how Artificial Intelligence is reshaping the human capabilities that occupations require. By analysing changes in abilities, skills, and knowledge between 2011 and 2025, we provide new evidence on how AI is transforming the human foundations of work.

Our findings reveal a clear dual pattern. Within occupation, capabilities that are more exposed to AI tend to experience larger increases in the level at which they are required within occupations. This suggests that AI reinforces a subset of abilities, skills, and knowledge that are complementary to the technology, raising the proficiency at which they are required within occupations.

However, the picture changes at the occupational level. Occupations with higher overall exposure to AI exhibit larger declines in the aggregate breadth of human capabilities required for work. Thus, while AI reinforces some capabilities, it is also associated with a compression of the broader capability profile of occupations.

These findings provide a new perspective on the relationship between technological change and deskilling. Our results suggest that an important part of the job transformation occurs at a deeper level, through changes in the capabilities that occupations require workers to possess. In this sense, our findings complement recent work emphasising the role of AI in reshaping expertise and task allocation within occupations (Autor and Thompson, 2025). They also provide a possible micro-foundation for broader arguments regarding the coexistence of substantial local adjustments and more limited aggregate effects of AI (Acemoglu, 2025). While AI may increase the required proficiency of selected capabilities, this does not necessarily translate into an expansion of the overall breadth of human capabilities required for work.

From a policy perspective, our findings suggest that the effects of AI exposure cannot be characterised in simple terms as either upskilling or deskilling. The same technological advances that increase the level of some capabilities may simultaneously reduce the relevance of others. Policies focused exclusively on developing AI-complementary capabilities may therefore overlook the broader risk of capability compression. Alongside investments in training and reskilling, policymakers should also consider how to preserve diverse and adaptable capability profiles that remain essential in rapidly changing labour markets.

Finally, our paper contributes methodologically by proposing a capability-based framework for measuring AI exposure. By exploiting observed improvements in AI benchmark performance, our approach provides a dynamic measure of technological progress that can be directly linked to changes in occupational capability requirements. Future research could extend this framework to other countries, occupations, and technological domains, as well as investigate whether the capability compression documented here translates into longer-run processes of occupational restructuring, deskilling, and labour-market inequality.

References

Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of Economic Literature*, 40(1):7–72.

Acemoglu, D. (2025). The simple macroeconomics of ai. *Economic Policy*, 40(121):13–58.

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4, chapter 12, pages 1043–1171. Elsevier, 1 edition.

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2020). AI and jobs: Evidence from online vacancies. NBER Working Paper 28257, National Bureau of Economic Research, Cambridge, MA.

Acemoglu, D. and Zilibotti, F. (2001). Productivity differences. *The Quarterly Journal of Economics*, 116(2):563–606.

Auer, R., Köpfer, D., and Sveda, J. (2024). The rise of generative AI: Modelling exposure, substitution and inequality effects on the us labour market. BIS Working Paper 1207, Bank for International Settlements (BIS).

Autor, D. and Thompson, N. (2025). Expertise. *Journal of the European Economic Association*, 23(4):1203–1271.

Autor, D. H. (2013). The “Task Approach” to labor markets: An overview. *Journal for Labour Market Research*, 46(3):185–199.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2006). The polarization of the u.s. labor market. *American Economic Review*, 96(2):189–194.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2008). Trends in u.s. wage inequality: Revising the revisionists. *Review of Economics and Statistics*, 90(2):300–323.

Autor, D. H., Katz, L. F., and Krueger, A. B. (1998). Computing inequality: Have computers changed the labor market? *The Quarterly Journal of Economics*, 113(4):1169–1213.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Beaudry, P., Green, D. A., and Sand, B. M. (2016). The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics*, 34(S1):199–247.

Bessen, J. (2018). AI and jobs: The role of demand. NBER Working Paper 24235, National Bureau of Economic Research, Cambridge, MA.

Borusyak, K., Hull, P., and Jaravel, X. (2025). A practical guide to shift-share instruments. *Journal of Economic Perspectives*, 39(1):181–204.

Braverman, H. (1974). Labor and monopoly capital. *Monthly Review*, 26(3):1–.

Brynjolfsson, E., Li, D., and Raymond, L. (2025). Generative ai at work. *The Quarterly Journal of Economics*, 140(2):889–942.

Brynjolfsson, E., Mitchell, T., and Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, 108:43–47.

Buchanan, B. G., Eckroth, J., and Smith, R. (2013). A virtual archive for the history of AI. *AI Magazine*, 34(2):86–86.

Eisfeldt, A. L., Schubert, G., Zhang, M. B., and Taska, B. (2023). Generative AI and firm values. Ssrn working paper, SSRN.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2023). Gpts are gpts: An early look at the labor market impact potential of large language models. arXiv preprint arXiv:2303.10130. Revised August 2023.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). Gpts are gpts: Labor market impact potential of llms. *Science*, 384:1306–1308.

Engberg, E., Koch, M., Lodefalk, M., and Schroeder, S. (2025). Artificial intelligence, tasks, skills, and wages: Worker-level evidence from germany. *Research Policy*, 54:105285.

Felten, E., Raj, M., and Seamans, R. (2021). Occupational, industry, and geographic exposure to Artificial Intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12):2195–2217.

Felten, E. W., Raj, M., and Seamans, R. (2019). The occupational impact of Artificial Intelligence: Labour, skills, and polarization. Ssrn scholarly paper, SSRN, Rochester, NY.

Fernández-Macías, E. and Hurley, J. (2017). Routine-biased technical change and job polarization in europe. *Socio-Economic Review*, 15(3):563–585.

Georgieff, A. and Hyee, R. (2021). Artificial Intelligence and employment: New cross-country evidence. OECD Social, Employment and Migration Working Papers 265, OECD Publishing.

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1):6.

Gmyrek, P., Berg, J., and Bescond, D. (2023). Generative AI and jobs: A global analysis of potential effects on job quantity and quality. ILO Working Paper 995324892702676, International Labour Organization (ILO).

Goldin, C. and Katz, L. F. (2007). The race between education and technology: The evolution of u.s. educational wage differentials, 1890 to 2005. NBER Working Paper 12984, National Bureau of Economic Research (NBER).

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Goos, M. and Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in britain. *The Review of Economics and Statistics*, 89(1):118–133.

Goos, M., Manning, A., and Salomons, A. (2009). Job polarization in europe. *American Economic Review*, 99(2):58–63.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, 104(8):2509–2526.

Hardy, W., Keister, R., and Lewandowski, P. (2018). Educational upgrading, structural change and the task composition of jobs in europe. *Economic Transition and Institutional Change*, 26(2):201–231.

Humlum, A. and Vestergaard, E. (2024). The adoption of ChatGPT. BFI Working Paper 2024-50, Becker Friedman Institute for Economics, University of Chicago.

Kapoor, R. and Teece, D. J. (2021). Three faces of technology's value creation: Emerging, enabling, embedding. *Strategy Science*, 6(1):1–4.

Katz, L. F. and Murphy, K. M. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *The Quarterly Journal of Economics*, 107(1):35–78.

Lassébie, J. and Quintini, G. (2022). What skills and abilities can automation technologies replicate and what does it mean for workers? new evidence. OECD Social, Employment and Migration Working Papers 282, OECD Publishing.

Manning, S. J. and Aguirre, T. (2026). How adaptable are american workers to ai-induced job displacement? NBER Working Paper 34705, National Bureau of Economic Research.

Marglin, S. A. (1974). What do bosses do? the origins and functions of hierarchy in capitalist production. *Review of Radical Political Economics*, 6(2):60–112.

Massenkoff, M. and McCrory, P. (2026). Labor market impacts of ai: A new measure and early evidence. Technical report, Anthropic.

McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., and Zolas, N. J. (2023). A.I. adoption in america: Who, what, and where. *Journal of Economics and Management Strategy*. Forthcoming.

Milanez, A. (2023). The impact of AI on the workplace: Evidence from oecd case studies of AI implementation. OECD Social, Employment and Migration Working Papers 289, OECD Publishing, Paris.

Montobbio, F., Staccioli, J., Virgillito, M. E., and Vivarelli, M. (2022). Robots and the origin of their labor-saving impact. *Technological Forecasting and Social Change*, 174:121122.

Montobbio, F., Staccioli, J., Virgillito, M. E., and Vivarelli, M. (2024). Labour-saving automation: A direct measure of occupational exposure. *The World Economy*, 47:332–361.

National Center for O*NET Development. O*net online help: Scales, ratings, and standardized scores. O*NET OnLine. Accessed: July 3, 2026.

National Center for O*NET Development. The O*NET® content model. O*NET Resource Center. s.d.; Recuperato il 14 maggio 2026, da `https://www.onetcenter.org/content.html`.

OECD (2023). *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*. OECD Publishing, Paris.

OECD (2025). Introducing the oecd ai capability indicators. Technical report, OECD Publishing, Paris.

OECD (2026). The oecd ai exposure measure: Mapping the oecd ai capability indicators to occupations. OECD Artificial Intelligence Papers 59, OECD Publishing, Paris.

Pizzinelli, C., Panton, A., Tavares, M. M., Cazzaniga, M., and Li, L. (2023). Labor market exposure to AI: Cross-country differences and distributional implications. IMF Working Paper 23/216, International Monetary Fund (IMF).

Reimers, N. and Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.

Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. *Journal of Labor Economics*, 24(2):235–270.

Taylor, F. W. (1911). *The Principles of Scientific Management*. Number taylor1911 in History of Economic Thought Books. McMaster University Archive for the History of Economic Thought.

Thompson, P. (1983). Braverman and the re-discovery of the labour process. In *The Nature of Work*. Palgrave, London.

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

Webb, M. (2019). The impact of Artificial Intelligence on the labor market. Ssrn working paper, SSRN.

Yin, C. and Zhang, Z. (2024). A study of sentence similarity based on the all-minilm-l6-v2 model with “same semantics, different structure” after fine tuning. In *Proceedings of the 2024 2nd International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2024)*, Advances in Computer Science Research. Atlantis Press.

A1 Human capabilities Described by O*NET

Table 14: Overview of O*NET Worker Requirements

Category	General Definition	Main Subdomains (Examples)
Abilities	Enduring attributes of the individual that influence performance.	*Cognitive Abilities* (verbal abilities, idea generation and reasoning, quantitative abilities, memory, perceptual and spatial abilities, attentiveness); *Psychomotor Abilities* (fine manipulation, control of movement, reaction time and speed); *Physical Abilities* (strength, endurance, flexibility, balance, coordination); *Sensory Abilities* (visual, auditory, and speech perception, including near/far vision, color discrimination, night vision, peripheral vision, depth perception, and speech recognition).
Skills	Developed capacities that facilitate learning or the rapid acquisition and application of knowledge.	*Content Skills* (reading comprehension, active listening, writing, speaking, mathematics, science); *Process Skills* (critical thinking, active learning, learning strategies, monitoring); *Social Skills* (social perceptiveness, coordination, persuasion, negotiation, instructing, service orientation); *Complex Problem-Solving Skills*; *Technical Skills* (operations analysis, technology design, programming, installation, operation and control, maintenance, troubleshooting, repairing, quality control analysis); *Systems Skills* (judgment and decision making, systems analysis, systems evaluation); *Resource Management Skills* (time, financial, material, and personnel resource management).

Continued on next page

Table 14 – continued from previous page

Category	General Definition	Main Subdomains (Examples)
Knowledge	Organized sets of principles and facts applying to broad content domains.	*Business and Management* (administration, economics, sales and marketing, human resources); *Manufacturing and Production* (production and processing, food production); *Engineering and Technology* (computers and electronics, engineering, design, construction, mechanical systems); *Mathematics and Science* (mathematics, physics, chemistry, biology, psychology, sociology, anthropology, geography); *Health Services* (medicine and dentistry, therapy and counseling, education and training); *Arts and Humanities* (languages, fine arts, history, archaeology, philosophy, theology); *Law and Public Safety* (public safety, security, law and government); *Communications and Transportation* (telecommunications, media, transportation); *Education* (required level of education, certifications, and subject-specific educational background).

Notes: Definitions are based on the O*NET Content Model (U.S. Department of Labor).

A2 Benchmarks Used in the AI Exposure Measure

Table 15: Core Benchmarks Used to Construct Ability, Skill, and Knowledge-Based AI Exposure Measures

Benchmark	**Description**
GSM8K	Grade-school reasoning problems requiring multi-step arithmetic reasoning.
MATH	12,500 competition-level mathematics problems with step-by-step solutions.
MATH-500	Hard subset of MATH, covering seven mathematical subjects.
AIME 2024	Olympiad-level reasoning benchmark requiring multi-step logical inference.
BBH	Big-Bench Hard: 23 challenging reasoning tasks beyond standard model capabilities.
HellaSwag	Adversarial commonsense reasoning dataset requiring physical and causal inference.
TruthfulQA	Measures whether models avoid common human misconceptions.
ARC-C	Challenge set of grade-school science questions requiring multi-hop reasoning.
ARC-E	Easier subset of ARC, solvable via factual recall with conceptual understanding.
MMLU	57-subject multitask exam spanning humanities, STEM, and social sciences.
MMLU-Pro	Harder variant with expanded options and filtered trivial questions.
HumanEval	Code synthesis benchmark measuring functional correctness of generated programs.
HumanEval+	Enhanced version with substantially expanded test coverage for robustness.
MBPP	974 Python programming tasks solvable by entry-level programmers.
MBPP+ / EvalPlus	MBPP enriched with substantially more test cases for rigorous evaluation.
SWE-bench Verified	Human-validated GitHub issues requiring multi-file code edits.
RepoBench	Repository-level code completion and retrieval tasks.
BigCodeBench	1,140 real-world programming tasks requiring tool and library usage.
MathVista	Multimodal mathematical reasoning benchmark combining figures, charts, and equations.

Continued on next page

Table 15 – continued from previous page

Benchmark	**Description**
ScienceQA	Multimodal science questions with diagrams, explanations, and reasoning chains.
VQA-v2	Visual question answering benchmark requiring recognition and reasoning.
Video-MME	Large-scale multimodal video understanding benchmark across 30 subfields.
LVBench	Long-video understanding benchmark with videos up to two hours.
RULER	Synthetic benchmark testing long-context capabilities with multi-hop tracing.
MRCR v2	Multi-round coreference resolution requiring tracking of multiple entities in long contexts.
LongBench v2	Long-context problem-solving benchmark across single- and multi-document reasoning.

Notes: The listed benchmarks are used to proxy AI capabilities across abilities, skills, and knowledge domains, forming the basis of the LLM-based AI exposure measure.

A3 Examples of Boolean Queries Used to Search AI-Related Publications

This appendix provides representative examples of the Boolean queries used to link O*NET worker requirements (Knowledge, Skills, and Abilities) to AI-related research publications. The queries combine semantic descriptors derived from O*NET definitions with AI-specific filters applied separately in the search process.

The examples below are illustrative and not exhaustive.

Knowledge

Engineering and Technology

```
("engineering principles" OR "engineering rules" OR "engineering knowledge" OR
 "engineering information" OR "engineering data" OR "engineering facts" OR
 "engineering concepts" OR "engineering theories" OR "engineering literature" OR
 "engineering book" OR "engineering books" OR "engineering publications" OR
 "engineering archive" OR "engineering archives" OR "engineering documentation" OR
 "engineering records" OR "engineering datasets" OR "engineering methods" OR
 "engineering techniques" OR "technology principles" OR "technology rules" OR
 "technology knowledge" OR "technology methods" OR "technology techniques")
```

Economics and Accounting

```
("economics principles" OR "economics rules" OR "economics knowledge" OR
 "economics information" OR "economics data" OR "economics facts" OR
 "economics concepts" OR "economics theories" OR "economics literature" OR
 "economics book" OR "economics books" OR "economics publications" OR
 "economics archive" OR "economics archives" OR "economics documentation" OR
 "economics records" OR "economics datasets" OR "economics methods" OR
 "economics techniques" OR "accounting principles" OR "accounting rules" OR
 "accounting knowledge" OR "accounting information" OR "accounting data" OR
 "accounting facts" OR "accounting concepts" OR "accounting theories" OR
 "accounting literature" OR "accounting book" OR "accounting books" OR
 "accounting publications" OR "accounting archive" OR "accounting archives" OR
 "accounting documentation" OR "accounting records" OR "accounting datasets" OR
 "accounting methods" OR "accounting techniques" OR "banking principles" OR
 "banking rules" OR "banking knowledge" OR "finance principles" OR
 "finance rules" OR "finance knowledge")
```

Philosophy and Theology

```
("philosophy and theology principles" OR "philosophy and theology rules" OR
 "philosophy and theology knowledge" OR "philosophy and theology information" OR
 "philosophy and theology data" OR "philosophy and theology facts" OR
 "philosophy and theology concepts" OR "philosophy and theology theories" OR
 "philosophy and theology literature" OR "philosophy and theology book" OR
 "philosophy and theology books" OR "philosophy and theology publications" OR
 "philosophy and theology archive" OR "philosophy and theology archives" OR
 "philosophy and theology documentation" OR "philosophy and theology records" OR
 "philosophy and theology datasets" OR "religious texts" OR
 "philosophical texts" OR "religious literature")
```

Arts and Humanities

```
("arts and humanities principles" OR "arts and humanities rules" OR
 "arts and humanities knowledge" OR "arts and humanities information" OR
 "arts and humanities data" OR "arts and humanities facts" OR
 "arts and humanities concepts" OR "arts and humanities theories" OR
 "arts and humanities literature" OR "arts and humanities book" OR
 "arts and humanities books" OR "arts and humanities publications" OR
 "arts and humanities archive" OR "arts and humanities archives" OR
 "arts and humanities documentation" OR "arts and humanities records" OR
 "arts and humanities datasets" OR "arts and humanities methods" OR
 "arts and humanities techniques")
```

Skills

Reading Comprehension *(Semantic component only; AI-specific filters are applied separately.)*

```
("reading comprehension" OR "text understanding" OR
 "natural language understanding" OR "document summarization" OR
 "question answering" OR "legal document analysis" OR
 "reading assistance" OR "understand sentences" OR
 "understand paragraphs")
```

Active Listening

```
("active listening" OR "conversational agents" OR
 "speech analysis" OR "listening comprehension" OR
 "dialogue systems" OR "customer service bots" OR
 "meeting transcription" OR "sentiment detection in speech" OR
 "give attention" OR "understand points" OR
 "ask questions" OR "avoid interruptions")
```

Critical Thinking

```
("decision support" OR "critical thinking models" OR
 "strategy evaluation" OR "policy analysis" OR
 "reasoning engines" OR "identify strengths" OR
 "identify weaknesses" OR "evaluate solutions")
```

Abilities

Manual Dexterity

```
("manual dexterity" OR
 ("move" AND "hands quickly") OR
 ("manipulate" AND "objects") OR
 "hand coordination")
```

Stamina

```
("stamina" OR
 ("exert" AND "physically") OR
 ("maintain" AND "physical activity") OR
 "sustained performance")
```

Memorization

```
("memorization" OR
 ("remember" AND "information") OR
 ("recall" AND "data") OR
 "memory encoding" OR
 "working memory")
```

Notes: These Boolean queries are illustrative examples. The full set of queries is generated programmatically from O*NET definitions and combined with AI-specific filters when querying publication databases.

A4 Top 10 Semantic Matches Between Benchmarks and human capabilities

As described in the Methodology section, we construct a semantic similarity measure linking textual descriptions of AI benchmarks to O*NET worker requirements (Abilities, Skills, and Knowledge). It is not a concern that certain worker requirements— such as Deductive Reasoning—appear multiple times among the top semantic matches. The semantic similarity procedure is used solely to identify conceptual links between benchmarks and O*NET requirements. As explained above, the AI exposure score for each worker requirement is ultimately computed as the average contribution of all benchmarks associated with it; therefore, repeated occurrences in the similarity ranking do not mechanically inflate its exposure measure.

Below we report the top semantic matches. For each pair, we show the benchmark name and description, together with the matched worker requirement and its definition.

Table 16: Top Semantic Similarity Links Between AI Benchmarks and O*NET Worker Requirements

Benchmark	**Benchmark Description**	**Matched Worker Requirement**
HMMT 2025	Harvard–MIT Mathematics Tournament 2025, a prestigious student-organized mathematics competition with individual tests, team rounds, and guts rounds held at MIT and Harvard.	**Deductive Reasoning** — The ability to apply general rules to specific problems to produce answers that make sense.
Natural2Code	NaturalCodeBench (NCB), a benchmark capturing complex real-world coding tasks across 402 problems in Python and Java derived from natural user queries.	**Idea Generation and Reasoning Abilities** — Abilities that influence the application and manipulation of information in problem solving.
Natural2Code	Same benchmark as above.	**Speech Recognition** — The ability to identify and understand the speech of another person.
Natural2Code	Same benchmark as above.	**Speed of Closure** — The ability to quickly make sense of, combine, and organize information into meaningful patterns.

Continued on next page

Table 16 – continued from previous page

Benchmark	**Benchmark Description**	**Matched Worker Requirement**
DS-FIM-Eval	DeepSeek's internal fill-in-the-middle evaluation dataset for assessing code completion performance in data science contexts.	**Deductive Reasoning** — The ability to apply general rules to specific problems to produce answers that make sense.
HumanEval+	Enhanced version of HumanEval with substantially expanded test coverage using EvalPlus.	**Idea Generation and Reasoning Abilities** — Abilities that influence the application and manipulation of information in problem solving.
WMT24++	Multilingual machine translation benchmark covering 55 languages and multiple content domains.	**Memorization** — The ability to remember information such as words, numbers, pictures, and procedures.
HMMT 2025	Same mathematics competition as above.	**Idea Generation and Reasoning Abilities** — Abilities that influence the application and manipulation of information in problem solving.
SWE-Bench Verified	Human-validated software engineering benchmark based on real GitHub issues.	**Speed of Closure** — The ability to quickly make sense of, combine, and organize information into meaningful patterns.
CodeForces	Competitive programming benchmark using algorithmic problems from the CodeForces platform.	**Deductive Reasoning** — The ability to apply general rules to specific problems to produce answers that make sense.

Notes: These matches reflect semantic similarity only and do not mechanically affect the exposure index of a given worker requirement, since AI exposure scores are computed as averages across all benchmarks linked to that requirement.

A5 Top Five and Bottom Five human capabilities by Exposure to the AI Topics Index

This appendix reports the worker requirements with the highest and lowest exposure to the AI Topics index. The top five worker requirements all display an exposure value equal to one, while the bottom five exhibit substantially lower exposure levels.

Table 17: Top Five and Bottom Five Worker Requirements by Exposure to the AI Topics Index

Worker Requirement	Exposure Value	Description
Top Five		
Active Learning (S)	1.00	Understanding the implications of new information for both current and future problem-solving and decision-making.
Arm–Hand Steadiness (A)	1.00	The ability to keep hand and arm steady while moving or holding them in one position.
Deductive Reasoning (A)	1.00	The ability to apply general rules to specific problems to produce logically coherent answers.
Mathematical Reasoning (A)	1.00	The ability to choose appropriate mathematical methods or formulas to solve a problem.
Judgment and Decision Making (S)	1.00	Considering the relative costs and benefits of potential actions to choose the most appropriate one.
Bottom Five		
Installation (S)	0.00	Installing equipment, machines, wiring, or programs to meet specifications.
Customer and Personal Services (K)	0.05	Knowledge of principles and processes for providing customer and personal services, including customer assessment and service quality evaluation.
Personnel and Human Resources (K)	0.07	Knowledge of principles and procedures for personnel recruitment, selection, training, compensation, and labor relations.

Continued on next page

Table 17 – continued from previous page

Worker Requirement	**Exposure Value**	**Description**
Dynamic Strength (A)	0.31	The ability to exert muscle force repeatedly or continuously over time.
Equipment Selection (S)	0.33	Determining the kind of tools and equipment needed to perform a job.

Notes: Letters in parentheses indicate the O*NET category of each worker requirement: (S) = Skill; (A) = Ability; (K) = Knowledge.

A6 Top Five and Bottom Five human capabilities by Exposure to the AI LLM Index

This appendix reports the worker requirements with the highest and lowest exposure to the AI LLM index. The top five worker requirements display relatively high exposure values, while the bottom five exhibit minimal or no exposure.

Table 18: Top Five and Bottom Five Worker Requirements by Exposure to the AI LLM Index

Worker Requirement	**Exposure Value**	**Description**
Top Five		
Deductive Reasoning (A)	1.00	The ability to apply general rules to specific problems to produce logically coherent answers.
Speed of Closure (A)	0.70	The ability to quickly make sense of, combine, and organize information into meaningful patterns.
Oral Comprehension (A)	0.57	The ability to listen to and understand information and ideas presented through spoken words and sentences.
Mathematical Reasoning (A)	0.56	The ability to choose appropriate mathematical methods or formulas to solve a problem.
Speech Recognition (A)	0.55	The ability to identify and understand the speech of another person.
Bottom Five		
Education and Training (K)	0.00	Knowledge of principles and methods for curriculum design, teaching, instruction, and training effectiveness.
Dynamic Strength (A)	0.001	The ability to exert muscle force repeatedly or continuously over time.
Sociology and Anthropology (K)	0.002	Knowledge of group behavior, societal trends, cultural differences, and human origins.

Continued on next page

Table 18 – continued from previous page

Worker Requirement	**Exposure Value**	**Description**
Public Safety and Security (K)	0.003	Knowledge of equipment, policies, procedures, and strategies for effective protection of people, data, and property.
Chemistry (K)	0.03	Knowledge of chemical composition, structure, properties, processes, and transformations of substances.

Notes: Letters in parentheses indicate the O*NET category of each worker requirement: (A) = Ability; (K) = Knowledge.

A7 Bartik-Style Shift-Share Instrument

A7.1 Motivation and methodology

A skeptical reader might argue that AI exposure could be endogenous if AI developments were driven by labour demand factors. AI companies could, in fact, anticipate demand for specific skills and focus their R&D efforts towards capabilities that are in high demand.

To appease these concerns, we adopt a Bartik IV (or Shift-Share) approach (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2025), taking advantage of variation in how a capability's pre-AI occupational footprint overlaps with the broader AI exposure of the occupations that use it. The intuition is that a capability is predicted to be more exposed when it is concentrated in occupations that are themselves heavily exposed to AI through their wider mix of capabilities. Because this predicted exposure is built from the pre-AI importance structure (which occupations relied on which capabilities before AI arrived) combined with subsequent, externally driven progress in AI capabilities, it is plausibly orthogonal to any contemporaneous response of AI development to labour demand.

The identifying assumption is that the pre-AI importance structure of occupations (which capabilities each occupation relied on, and how capabilities were distributed across occupations) is predetermined with respect to subsequent AI progress, so that predicted exposure shifts a capability's AI exposure without itself being driven by latent demand for AI.

The Bartik is defined at the capability level c as follows:

$$Bartik_c = \sum_o Share_{oc,PRE} \times AiExposure_{o,POST}, \tag{1}$$

where $Share_{oc,PRE}$ is the share of capability c's total importance accounted for by occupation o in the pre-AI period (the proportion of a capability located within an occupation):

$$Share_{oc,PRE} = \frac{Importance_{oc,PRE}}{\sum_{o'} Importance_{o'c,PRE}}. \tag{2}$$

The term $AiExposure_{o,POST}$ is the overall exposure of occupation o to the AI shock, obtained by distributing the capability-level shock across occupations according to each capability's importance within the occupation:

$$AiExposure_{o,POST} = \sum_c AiExposure_{c,POST} \times \frac{Importance_{oc,PRE}}{\sum_{c'} Importance_{oc',PRE}}. \tag{3}$$

The *Share* measure captures the proportion of a capability located within an occupation (summing to one across occupations, for a fixed capability), whereas the exposure term distributes the capability-level shock across occupations by weighting it by the proportion of each capability within an occupation (summing to one across capabilities, for a fixed occupation). These are distinct objects.

A7.2 Results

Results from the Bartik estimations are provided below in Table 19. We find in columns (1) and (3) that the capability-wide instruments strongly correlate with the corresponding occupation–capability AI exposure measures, both for the LLM Stats-based and the AI Topics-based exposure. This leads to strong first stages, with F-statistics well above conventional threshold levels.

The second-stage estimates (columns 2 and 4) are statistically significant and support our main results. The estimated effects are 0.042 for AI LLM and 0.033 for AI Topics, which are slightly larger than or broadly consistent with our main estimates, as expected from IV estimation.

We then conclude that AI exposure is plausibly exogenous.

Table 19: **Bartik IV Estimates**

	AI LLM		**AI Topics**	
	(1) First Stage	(2) Second Stage	(3) First Stage	(4) Second Stage
Bartik Capability	1.246***		1.959***	
	(0.013)		(0.019)	
AI Exposure		0.042***		0.033***
		(0.004)		(0.005)
First-stage F-statistic	8,625.86		10,695.66	
Observations	78,117	78,117	74,147	74,147

Notes: The table reports first-stage and second-stage estimates from Bartik-style instrumental variable specifications. The first-stage columns report the relationship between the Bartik instrument and the endogenous AI exposure measures, while the second-stage columns report the estimated effect of instrumented AI exposure on changes in capability requirements. Standard errors are reported in parentheses. The first-stage F-statistic is reported as a measure of instrument strength. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.