From Exposure to Adoption: AI in Colombia's Labor Market Using Online Vacancies and Household Surveys

Andrés Garcia-Suaza* Alexander Sarango-Iturralde†
Pamela Caiza-Guamán‡ Mateo Gil Díaz§ Dana Acosta Castillo¶

February 3, 2026

Abstract

The rapid advancements in the domain of artificial intelligence (AI) have exerted a considerable influence on the labor market, thereby engendering alterations in the demand for specific skills and the structure of employment. This study aims to evaluate the extent of exposure to AI within the Colombian labor market and its relation with workforce characteristics and available job openings. To this end, we built a specific AI exposure index for Colombia based on skill demand in job posts. Our findings indicate that 33.8% of workers are highly exposed to AI, with variations observed depending on the measurement method employed. Furthermore, it is revealed a positive and significant correlation between AI exposure and wages, i.e., highly exposed to AI earn a wage premium of 21.8%. On the demand side, only 2.5% of job openings explicitly mention AI-related skills. These findings imply that international indices may underestimate the wage premium associated with AI exposure in Colombia and underscore the potential unequal effects on wages distribution among different demographic groups.

Keywords: Artificial intelligence; labor market; job posts; occupations; skills; Colombia.
JEL Codes: E24; J23; J24; O33.

*Associate professor, School of economics at Universidad del Rosario. Research fellow Global Labor Organization (GLO). Email: andres.garcia@urosario.edu.co.

†PhD candidate, Université Paris 1, Panthéon–Sorbonne & Institut de Recherche pour le Développement (IRD), UMR Développement et Sociétés (DEVSOC). I acknowledge the support from the IRD through a doctoral fellowship at DEVSOC. Email: alexandersarangoiturralde@outlook.com.

‡Research Assistant, Universidad del Rosario, Email: raiza.caiza@urosario.edu.co.

§Research Assistant, Universidad del Rosario, Email: mateo.gil@urosario.edu.co.

¶Research Assistant, Universidad del Rosario, Email: dana.acosta@urosario.edu.co.

We would like to express our gratitude to The Future of Life Institute for funding this paper through the grant titled "Evaluating the Impact of AI on Poverty, Health, Energy, and Climate SDGs". We thank to Carlo Pizzinelli and his co-authors for sharing with us their complementarity index and C-AIOE estimates.

1 Introduction

Artificial intelligence (AI) is reshaping production processes and the task content of jobs, with implications for productivity and labor demand (Georgieff and Hyee, 2022). Recent evidence shows that AI can raise worker productivity and alter task allocation, although effects are heterogeneous across jobs, firms, and workers (Carbonero et al., 2023; Noy and Zhang, 2023). Because diffusion and complementary inputs (human capital, organizational capabilities, and infrastructure) are uneven across countries, the labor-market implications of AI are likely to differ markedly across developing and advanced economies.

We study how potential occupational exposure to AI compares with revealed AI-related capability demand in hiring, and how both relate to wages and workforce composition in a high-informality labor market. There are generally competing forces influencing the costs and benefits of AI. On one hand, AI can automate some tasks and augment others, including non-routine cognitive tasks. As a result, potential exposure reflects technological applicability rather than an unambiguous automation risk; displacement risk is more likely where task bundles are routinizable and codifiable (Brynjolfsson et al., 2018; Acemoglu and Restrepo, 2019; Green, 2024). In contrast, jobs that rely on cognitive skills, such as critical thinking, creativity, and problem solving, tend to be less impacted by automation (Autor and Dorn, 2013). These occupations may also experience productivity increases, particularly in sectors that integrate advanced technologies such as agritech or fintech (see, for example, Mhlanga, 2021; Alonso et al., 2022). Georgieff and Hyee (2022) and Alonso et al. (2022) show that occupations involving computer use tend to grow more rapidly, while occupations with low computer use experience a reduction in the average hours worked. Moreover, AI adoption can generate new task bundles and new job profiles, particularly in digital-intensive activities. Acemoglu and Restrepo (2019) and Carbonero et al. (2023) argue that there is a sort of reintegration effect, in which the introduction of new AI-related tasks creates employment growth opportunities, even as automation leads job losses in other sectors.

The labor market characteristics of developing countries, such as those in Latin America, highlight the importance of studying the impacts of AI. Specifically, developing countries are characterized by high levels of informality, which correlates with lower exposure to AI (Cazzaniga et al., 2024). Further, informal workers are less educated and face significant human capital misallocation. Additionally, low levels of technological investment by firms in developing countries may prolong AI adoption (Fontanelli et al., 2025). Lower investment and weaker complementary capabilities can slow adoption and change the incidence of AI-related gains across workers and firms. Nonetheless, some evidence suggests that digital skill acquisition has accelerated in response to technological

changes following the COVID-19 pandemic. For Colombia, Sarango-Iturralde et al. (2021) reports a 3 p.p. increase in the demand for digital competencies following the pandemic, consistent with accelerated skill adjustment in response to technological change.

This study examines AI-related exposure and adoption patterns from both the labor-demand and labor-supply sides in Colombia. On the demand side, we use online job postings to construct two vacancy-based measures of AI-related capability demand. First, we measure the share of postings that explicitly mention AI-related skills. Second, we build AI-Col, a continuous vacancy-level measure of AI-related capability demand based on an AI-term dictionary and TF-IDF weighting, which captures the intensity of AI-related terminology beyond explicit mentions. We then aggregate vacancy-level measures to the occupation level, making them directly comparable to standard occupational exposure indices, and map these occupation-level measures to workers in household surveys. On the supply side, we use occupation-level exposure indices proposed by Webb (2020), Felten et al. (2021), and Pizzinelli et al. (2023) to document how exposure varies with education, wages, and worker demographics in household surveys.

Recent cross-country work operationalizes this heterogeneity by jointly classifying occupations by exposure and complementarity. For example, Cazzaniga et al. (2024) classifies occupations into three categories: high exposure and high complementarity, high exposure and low complementarity, and low exposure. The first category is likely to have increased productivity with a low risk of displacement, including occupations such as surgeons, lawyers, and judges. By contrast, the second category faces a higher risk of automation (Su et al., 2021). These types of occupation, e.g. call center operators, may be more likely to experience reallocation across employment arrangements, potentially including formal and informal jobs. Gmyrek et al. (2024) conduct a similar cross-country study that categorizes occupations into three groups: those with potential for automation, those with potential for growth, and a final group termed *Big unknown* where the effects of AI remain uncertain. These classifications are informative, but they are typically silent on how adoption signals appear in labor demand, particularly in high-informality settings.

This paper contributes in three ways. First, it provides new evidence on AI-related patterns in a Global South labor market characterized by high informality and uneven diffusion. Second, it benchmarks Colombia using three widely used exposure indices and shows how conclusions depend on conceptual emphasis and data sources. Third, it introduces AI-Col, a vacancy-based measure of revealed AI-related capability demand that we aggregate to occupations and map to workers, enabling a direct comparison between potential exposure and revealed adoption at the hiring margin.

In terms of labor demand, our analysis of over one million job postings reveals

that only 2.5% explicitly mention AI-related skills, a figure that is lower than in developed economies (see, for example, Acemoglu et al., 2022). The terminology used in Colombian postings is consistent with applied AI and with broader digital competencies that support task performance, rather than narrow, specialized AI roles. Moreover, AI-related job postings are associated with 11% higher posted wages. However, the share of vacancies that explicitly mention AI skills is sensitive to how employers label requirements. In Colombia, AI-related capabilities are frequently framed within broader digital competencies, which motivates complementing the binary "AI mention" measure with a continuous AI-intensity index (AI-Col).

We then translate these vacancy-level indicators into occupation-level measures and map them to workers in household surveys. We estimate that 33.8% of workers are employed in occupations with high average AI-Col intensity. Regarding wage differences, workers in high-AI-Col occupations receiving a 21.8% higher wages on average, consistent with estimates from advanced economies. Oaxaca-Blinder decompositions indicate that this premium is only partially explained by observed characteristics, suggesting that AI-related skills are valued in the labor market beyond traditional qualifications.

Taken together, these demand- and supply-side patterns raise a measurement question: how should we interpret occupation-based AI exposure indices relative to vacancy-based measures of revealed AI-related capability demand? The next subsection clarifies what each measure captures and why they need not coincide.

1.1 From occupational exposure to revealed adoption: what each measure captures

A growing literature develops occupation-level measures of AI exposure that quantify the overlap between AI capabilities and occupational content (tasks, skills, or abilities). In this view, exposure captures technological applicability, defined as the extent to which AI systems can perform or assist with activities embedded in occupational content, rather than a direct prediction of labor-market outcomes (Manning and Aguirre, 2026).

Importantly, exposure does not mechanically translate into either substitution or complementarity between AI and labor. Some influential contributions emphasize AI's potential to replace human labor and are therefore closer to measures of automation risk (Frey and Osborne, 2017; Eloundou et al., 2024; Brynjolfsson et al., 2018; Webb, 2020). At the same time, other strands of the literature stress that AI can also augment workers by enhancing productivity and reorganizing tasks rather than fully displacing labor, so exposure is not a sufficient statistic for the sign or magnitude of employment and wage effects (Felten et al., 2021; Prytkova et al., 2024). This ambiguity is consistent with the

view that partial automation of some tasks can coexist with the creation of new tasks and new forms of work organization (e.g., Brynjolfsson et al., 2018).

Within this family of exposure measures, we rely on three widely used indices that differ in both data inputs and conceptual emphasis. Felten et al. (2021) construct the AI Occupational Exposure (AIOE) index by mapping a set of AI applications to occupational abilities, remaining agnostic about whether AI ultimately substitutes for or complements those abilities. Pizzinelli et al. (2023) build on the AIOE by introducing a potential-complementarity adjustment that downweights exposure in occupations where AI is more likely to operate as a complementary technology; the resulting complementarity-adjusted measure (C-AIOE) therefore assigns higher values to occupations that combine high exposure with limited complementarity. In a distinct approach, Webb (2020) measure exposure by quantifying textual overlap between AI-related patents and job descriptions. Taken together, these indices provide a coherent benchmark of potential exposure across occupations while differing in how they map AI technologies to occupational content and in whether complementarity is explicitly incorporated.

While exposure indices are informative about potential technological applicability, they do not directly capture whether firms adopt AI technologies and reorganize production accordingly, nor the organizational and skill-related prerequisites required for deployment. Firm-level data on AI adoption are scarce, and a growing body of work therefore uses technology-related human capital as a proxy for adoption, leveraging the detailed text in online vacancies to observe explicit demand for AI-related skills (Alekseeva et al., 2021; Calvino et al., 2022). Copestake et al. (2024) proxy firm-level adoption using AI-skill requirements in job adverts, while emphasizing that online vacancies are not representative of all hiring and therefore capture only a subset of firms' recruiting behavior. This perspective is consistent with evidence showing that AI-related vacancy activity is systematically related to firms' occupational structures and exposure measures, suggesting that hiring-based measures can capture meaningful variation in realized AI activity (Acemoglu et al., 2022). At the same time, vacancy-based measures primarily capture adoption at the hiring margin, and may understate adoption occurring through internal upskilling or vendor solutions without new recruiting (Alekseeva et al., 2021; Calvino et al., 2022).

Building on this conceptual separation, our empirical strategy integrates both perspectives. We contrast internationally used occupation-level exposure measures (Webb, 2020; Felten et al., 2021; Pizzinelli et al., 2023) with a country-specific index derived from job-posting texts (AI-Col), which captures firms' revealed demand for AI-related capabilities in Colombia. This design characterizes potential exposure on

the labor-supply side while documenting the penetration of AI-related skill demand on the labor-demand side, recognizing that these measures capture distinct objects. The imputed occupation codes also allow us to assign internationally used exposure indices to the same vacancy data, facilitating a like-for-like comparison between potential exposure (task/ability-based) and revealed capability demand at the hiring margin.

This paper is organized as follows. Section 2 describes the data and empirical approach. Sections 3 and 4 present the results, combining evidence from online job postings (labor demand) and household survey microdata (labor supply) to characterize revealed AI-related capability demand and potential occupational exposure. Section 5 concludes.

2 Data and methodology

Our empirical design combines (i) a large dataset of online job postings to measure employers' revealed demand for AI-related capabilities at the hiring margin, and (ii) nationally representative household survey microdata to describe workers' occupational allocation and outcomes. We benchmark vacancy-based hiring signals against three widely used occupation-based indices that capture potential AI applicability in occupational content (tasks/abilities). Throughout, we present descriptive patterns and do not interpret estimated associations as causal, consistent with the view that hiring data and exposure indices capture different objects and have distinct coverage limitations.

Online job postings (labor demand). We collect 1,169,943 online job ads posted between March 2022 and December 2023 from main job-portals, one of the largest job boards in Colombia. For each ad we observe the job title, full text description, required education, and (when available) posted wages. We drop duplicate postings (identical title–text pairs) and ads with insufficient text to support occupational coding. Online postings provide a direct window into skill requirements at the hiring margin, but they do not represent the universe of hiring—especially in high-informality settings—and tend to over-represent more formal and higher-skill segments of labor demand. We therefore interpret vacancy-based measures as posted-vacancy signals and focus on within-sample comparisons across occupations and requirements rather than on population-level adoption rates.

Household survey (labor supply and outcomes). We use the 2023 wave of the Gran Encuesta Integrada de Hogares (GEIH), Colombia's nationally representative household survey. The GEIH provides information on wages, education, demographics, sector, and occupations coded in the four-digit Clasificación Única de Ocupaciones para Colombia (CUOC), a national adaptation aligned with ISCO-08 (22,520,713 employees).

2.1 Occupational coding and crosswalks

The postings dataset does not provide occupational codes. To harmonize vacancies with worker data and with occupation-based exposure indices, we impute an occupation code for each posting using a text-based classifier that maps vacancy titles and descriptions into the CUOC taxonomy. The classifier adapts the TF–IDF retrieval approach implemented in the `labourR` package (originally designed for ESCO) to the Colombian CUOC classification (Kouretsis et al., 2020; Garcia-Suaza et al., 2025). Because automated coding from free text is imperfect—especially at highly detailed levels—we emphasize analyses at aggregated occupational levels (3-digit ISCO-08) when linking vacancies to workers and when benchmarking exposure indices, which mitigates misclassification noise that would otherwise attenuate estimated relationships.

Exposure indices and Colombian data sources do not coincide in occupational taxonomies. For Felten et al.'s AIOE and Pizzinelli et al.'s C-AIOE (defined in O*NET-SOC/SOC), we map SOC into ISCO-08 and then into CUOC using official crosswalks.[1] For Webb's measure (defined in Census OCC-1990), we map OCC-1990 into ISCO-88 and then into ISCO-08 before aligning to CUOC. Because these crosswalks are many-to-many, we implement a transparent uniform-splitting rule to avoid double counting: when one source code maps to multiple destination codes, we divide the originating index value equally across mapped destinations and then aggregate to the target level. We use this rule as a reproducible approximation and report results at aggregated occupational levels where assignments are more stable.

2.2 Measures: potential exposure vs. revealed capability demand

We distinguish (i) potential occupational exposure—technological applicability based on task/ability content—from (ii) revealed AI-related capability demand observed in vacancy texts at the hiring margin. These objects need not coincide.

[1] For SOC–ISCO crosswalks, see the U.S. Bureau of Labor Statistics tables: `https://www.bls.gov/soc/isco_soc_crosswalk.xls`.

We use three widely used indices that map AI technologies to occupational content: Felten et al. (2021) construct the AI Occupational Exposure index (AIOE) by linking AI applications to O*NET abilities; Pizzinelli et al. (2023) adjust AIOE for potential complementarity (C-AIOE); and Webb (2020) measure exposure using textual overlap between AI-related patents and job descriptions. Details are provided in Appendix A.

From vacancy texts, we build two labor-demand indicators: (i) an explicit AI mention dummy based on a narrow set of core AI terms, and (ii) AI-Col, a continuous index that captures broader AI-related capability requirements even when employers bundle them under general digital skill language.

2.3 Construction of AI-Col (vacancy-level)

AI-Col is a vacancy-level index of revealed AI-related capability demand. It is not intended to measure economy-wide "AI exposure" or automation risk for all occupations. Instead, it captures the extent to which postings explicitly require AI-related capabilities in recruitment language—and is therefore closer to AI-specialist / AI-using hiring signals than to broad occupational exposure. This distinction addresses concerns that vacancy-based AI terminology may primarily proxy demand for AI technical profiles rather than task-level exposure across all jobs.

Dictionary and text processing. We apply a curated bilingual (Spanish/English) dictionary of AI-related terms to vacancy titles and descriptions after standard preprocessing (lowercasing, punctuation removal, stop-word removal, and tokenization). The dictionary is anchored in established vacancy- and text-based approaches to measuring AI-related skill demand (Alekseeva et al., 2021; Babashahi et al., 2024; Baruffaldi et al., 2020; Dawson et al., 2021; Dernis et al., 2023; Gehlhaus, 2021; Squicciarini and Nachtigall, 2021) and is expanded using practitioner taxonomies to improve coverage of emerging terminology (e.g., AI-use and ecosystem taxonomies) (Theofanos et al., 2024). We manually reviewed and curated candidate terms to minimize false positives arising from generic business or IT language and from context-dependent abbreviations (Appendix Table C.1 reports term families and illustrative examples; the full dictionary is provided in the supplementary materials / online repository). To avoid mechanically assigning higher scores to longer ads, we treat term frequency as binary (presence/absence).

IDF-weighted scoring. AI-Col aggregates matched terms using inverse document frequency (IDF) weights to emphasize terms that are more diagnostic of AI-related

content and to reduce noise from generic digital terminology. The weighting scheme is a signal-extraction device (discriminative power), not a measure of economic importance of terms.

$$IDF(t) = \log\left(1 + \frac{T}{v(t)+1}\right),$$

where T is the total number of vacancies and $v(t)$ is the number of vacancies in which term t appears. The AI-Col score for vacancy j is

$$AI\text{-}Col_j = \sum_{t \in \mathcal{D} \cap \text{text}_j} IDF(t),$$

where $\mathcal{D}$ denotes the AI-term dictionary and text_j is the tokenized text (title + description) of vacancy j.

We match dictionary terms allowing minor spelling variation using Jaro–Winkler string distance and classify a match when distance ≤ 0.05 (i.e., similarity ≥ 0.95). Vacancies with no detected dictionary terms receive $AI\text{-}Col_j = 0$.

Operationalization in labor-demand and labor-supply analyses. Given the skewness of AI-Col, we use it in two complementary ways. On the labor-demand side, we use vacancy-level AI-Col directly and also report the explicit AI-mention rate based on core terms. On the labor-supply side, we aggregate AI-Col to the 3-digit ISCO-08 level by averaging across vacancies within each occupational group, and we use the median of this occupation-level distribution as a relative threshold to classify occupations as higher- vs. lower-intensity in AI-related capability demand (following the median-split approach used in related work, e.g. Cazzaniga et al., 2024). This classification is intended for within-sample ranking rather than as a structural parameter of exposure.

3 AI-related capability demand and occupational exposure: integrated evidence from job postings and household data

This section combines two complementary measurement lenses. We start from labor demand, using online vacancies to quantify employers' revealed demand for AI-related capabilities at the hiring margin. We then link these hiring signals to worker allocation and wage patterns in household microdata. Throughout, vacancy-based measures are interpreted as signals among firms and jobs that recruit online, while occupation-based indices capture potential AI applicability embedded in occupational task and ability

content.

A central empirical implication is that potential exposure and revealed capability demand need not coincide. Potential exposure indices assign a fixed score to an occupation and apply it uniformly to all jobs mapped into that occupation. Vacancy-based measures, by contrast, vary within occupations because they are constructed from the text of each posting. The difference is not a contradiction; it is informative about diffusion, labeling practices, and heterogeneity within occupational titles.

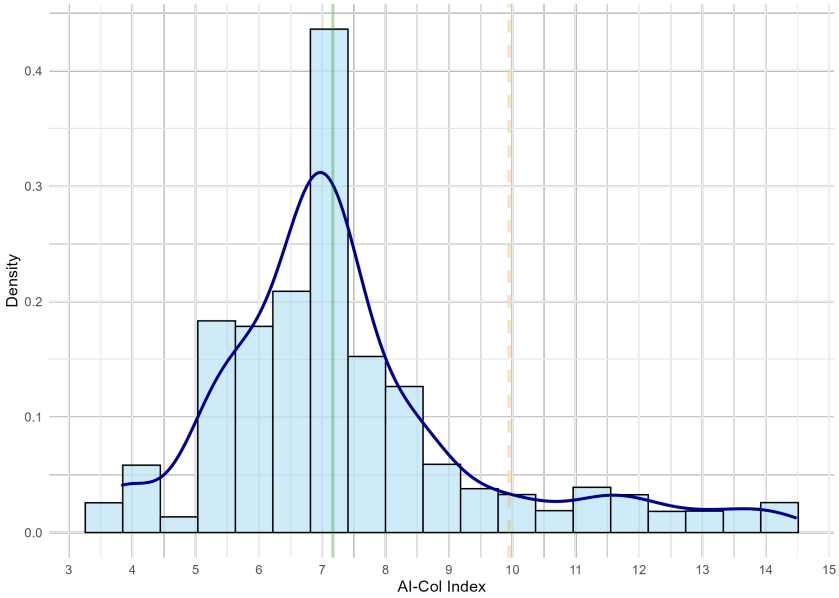
We document four sets of results. First, we establish prevalence levels for AI-related capability demand in vacancies and compare them to the prevalence implied by occupation-based exposure indices. Second, we show how both vacancy-based and occupation-based measures concentrate in specific occupational groups. Third, we relate exposure measures to wages, using posted wages in vacancies and realized wages in the workforce. Finally, we use decomposition evidence to clarify which observed worker characteristics account for wage gaps between high- and low-exposure jobs and how much remains as residual differentials.

Online postings provide unusually rich information on skill requirements at the hiring margin, but they do not represent the universe of hiring in a high-informality economy. They tend to over-represent more formal, higher-skill segments of labor demand. For that reason, we emphasize within-sample comparisons across occupations and requirements, and we interpret levels, such as the share of ads mentioning AI, with caution.

How prevalent is AI-related capability demand in posted vacancies? We begin with a prevalence question: how often do employers explicitly or implicitly request AI-related capabilities in posted vacancies? Using the vacancy-level AI-Col measure, only a small fraction of postings contain any AI-related term. This low prevalence provides a useful benchmark: even when potential exposure indices suggest broad technological applicability, the hiring margin may reveal limited explicit adoption signals in job advertisements. The gap between these objects is one of the key empirical patterns in the paper.

To characterize the intensity margin, Figure 1 plots the distribution of the AI-Col index across vacancies. The distribution is strongly right-skewed, with mass at zero and a long right tail among the exposed subset. This motivates reporting both a binary indicator for any AI-related term and a continuous intensity score, and it motivates interpreting AI-Col primarily as a ranking device within the vacancy corpus rather than as a structural parameter of exposure.

Figure 1: Distribution of the AI-Col Index across job vacancies



Notes: The green solid line represents the median of the distribution, while the orange dashed line marks the mean. Outlier observations were excluded from the graph. Source: Authors' calculations based on job postings database (2022, 2023).

How does vacancy-based demand compare to workforce exposure patterns? We anchor the demand evidence in worker-side exposure measured in the household survey. Table 1 reports the distribution of workers across low- and high-exposure occupations in the GEIH. For comparability across indices, we define high exposure as values above the median of the respective exposure distribution. The occupation-based indices imply that roughly two-fifths of workers are in high-exposure occupations, which is consistent with patterns reported for other middle-income settings. Importantly, AI-Col is constructed from vacancies and is then aggregated to an occupational level for worker-side comparisons; under the median-split classification, AI-Col assigns a lower high-exposure share than the occupation-based indices. This difference is informative: it is consistent with a setting in which AI-related capability demand is concentrated in a narrower set of jobs than what potential exposure metrics would suggest.

Table 1: Distribution of workers by level of AI exposure

Level	Quantity				Percentages			
	Webb (2020)	**Felten** *et al.* (2021)	**Pizzinelli** *et al.* (2023)	**AI-Col** Index	**Webb** (2020)	**Felten** *et al.* (2021)	**Pizzinelli** *et al.* (2023)	**AI-Col** Index
Low exposure	13,693,453	13,414,802	13,171,105	14,904,285	60.80%	59.60%	58.50%	66.20%
High exposure	8,827,260	9,105,911	9,349,608	7,616,428	39.20%	40.40%	41.50%	33.80%
Total	22,520,713	22,520,713	22,520,713	22,520,713	100.00%	100.00%	100.00%	100.00%

Notes: All calculations were performed using sample weights. The occupations categorized as "High Exposure" have values exceeding the median of their respective exposure index. Source: Authors' calculations based on GEIH (2023).

A useful way to interpret cross-index differences is to examine overlap in occupational classifications. Because indices draw on different inputs and emphasize different channels, they can agree in aggregate while disagreeing in the identity of high-exposure occupations. Table 2 illustrates this point by cross-tabulating Webb's measure, which emphasizes automation potential through patent-to-text overlap, and Pizzinelli et al.'s complementarity-adjusted measure, which is designed to capture the scope for augmentation as well as automation. Occupations that both indices label as high exposure include managerial and professional groups that rely on analytical tasks and structured information processing. Occupations that both indices label as low exposure tend to involve manual, operational, or face-to-face service content. The informative cases are the disagreements: some high-status occupations are classified as more exposed under complementarity-oriented measures because AI can augment cognitive and decision-making tasks, even if full automation risk is limited. This contrast helps clarify what each index is capturing.

Table 2: Comparison of AI Exposure Indices: Coincidences between Webb (2020) and Pizzinelli et al. (2023)

	Pizzinelli *et al.* (2023) index	
Webb (2020) index	**High Exposure**	**Low Exposure**
High Exposure	• Business Services and Administration • Sales, Marketing and Development • Finance Professionals • Legal Professionals • Medical and Pharmaceutical Technicians	• Production Managers in Agriculture • Medical Doctors • Mining, Manufacturing and Construction • Subsistence Crop Farmers • Manufacturing Laborers
Low Exposure	• Legislators and Senior Officials • Managing Directors and Chief Executives • University and Higher Education Teachers • Creative and Performing Artists • Secretaries (general)	• Hotel and Restaurant Managers • Sports and Fitness Workers • Personal Care Workers in Health Services • Machinery Mechanics and Repairers • Food Processing and Related Trades Workers

Notes: Source: Prepared by the authors. The occupations categorized as "High Exposure" have values exceeding the median in the Pizzinelli et al (2023) and Webb (2020) indices, respectively.

Worker sorting by exposure: demographics, education, and informality. Table 3 compares worker characteristics across low- and high-exposure occupations. The patterns are consistent with occupational sorting: more exposed occupations tend to be staffed by more educated workers and are more likely to be formal and salaried. In a high-informality economy, this distinction is salient. Across several indices, high-exposure jobs are disproportionately formal, suggesting that AI-related capabilities are requested and rewarded primarily in segments with higher market access, stronger organizational capacity, and greater use of standardized recruitment channels. Gender composition differs across indices, which is consistent with the idea that measures emphasizing automation potential and measures emphasizing augmentation can map differently into gendered occupational segregation. Age gradients are comparatively modest.

Table 3: Characteristics of workers by AI exposure level

Variable / category	Webb (2020)		Felten *et al.* (2021)		Pizzinelli *et al.* (2023)		AI-Col Index	
	Low exposure	High exposure	Low exposure	High exposure	Low exposure	High exposure	Low exposure	High exposure
Age More than 30	73.94%	75.77%	74.46%	74.94%	75.86%	72.95%	73.76%	76.41%
Age Less than 30	26.06%	24.23%	25.54%	25.06%	24.14%	27.05%	26.24%	23.59%
Sex Men	50.77%	71.46%	68.01%	45.44%	69.35%	44.13%	60.71%	55.31%
Sex Women	49.23%	28.54%	31.99%	54.56%	30.65%	55.87%	39.29%	44.69%
Education No University	83.11%	71.99%	95.17%	54.55%	92.60%	59.24%	93.40%	50.09%
Education University	16.89%	28.01%	4.83%	45.45%	7.40%	40.76%	6.60%	49.91%
Informality Informal	54.79%	58.02%	69.70%	35.97%	67.56%	39.86%	67.01%	34.63%
Informality Formal	45.21%	41.98%	30.30%	64.03%	32.44%	60.14%	32.99%	65.37%
Employment status Salaried workers	51.78%	41.61%	39.20%	60.45%	37.80%	61.87%	39.90%	63.25%
Employment status Self-employed / Employed	41.26%	48.68%	49.05%	36.97%	50.41%	35.37%	48.40%	35.89%
Employment status Others	6.97%	9.71%	11.75%	2.58%	11.79%	2.75%	11.71%	0.87%

Notes: Source: Authors' calculations based on GEIH (2023). All columns sum to 100 percent, and occupations categorized as "High Exposure" have values exceeding the median of their respective exposure index.

The occupational-group breakdown provides an additional consistency check. Table 4 shows that exposure assignments differ sharply across indices at the 1-digit ISCO level. Webb's measure allocates more high-exposure employment to routine-intensive, lower-wage groups relative to the other indices, which is consistent with its automation-oriented construction. Felten and Pizzinelli, by contrast, allocate high exposure more heavily to professional and technical groups. AI-Col, when mapped into the workforce at an occupational level, places the bulk of high-exposure employment in Professionals and related high-skill categories. These contrasts reinforce that cross-index comparisons are not primarily about statistical noise; they reflect different conceptual objects.

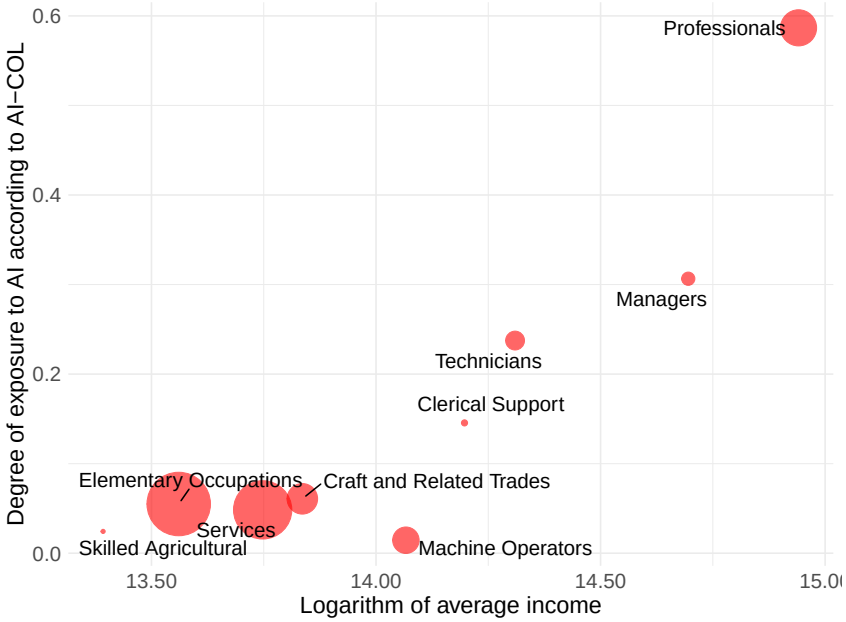
Table 4: Employment share by occupational groups and level of exposure to AI

Occupation (ISCO-08 major group)	Webb (2020)		Felten *et al.* (2021)		Pizzinelli *et al.* (2023)		AI-Col Index		Total
	Low exposure	High exposure	Low exposure	High exposure	Low exposure	High exposure	Low exposure	High exposure	Share
Managers	2.08%	12.99%	0.94%	14.33%	5.07%	8.17%	0.00%	18.80%	6.36%
Professionals	6.36%	18.81%	0.64%	26.86%	2.61%	23.41%	0.41%	32.44%	11.24%
Technicians	5.23%	9.89%	2.94%	13.11%	3.10%	12.63%	1.01%	18.89%	7.05%
Clerical Support	9.78%	0.00%	0.00%	14.71%	0.00%	14.33%	4.26%	9.25%	5.95%
Services and Sales	34.81%	0.00%	16.14%	28.57%	12.79%	32.97%	31.72%	0.52%	21.17%
Skilled Agricultural	1.00%	13.58%	9.96%	0.00%	10.14%	0.00%	8.96%	0.00%	5.93%
Craft and Related Trades	10.54%	8.01%	16.03%	0.00%	13.78%	3.58%	11.69%	5.35%	9.55%
Machine Operators	12.09%	2.96%	14.29%	0.00%	12.74%	2.56%	12.87%	0.00%	8.51%
Elementary Occupations	18.11%	33.75%	39.05%	2.42%	39.78%	2.36%	29.09%	14.76%	24.24%

Notes: Source: Authors' calculations based on GEIH (2023). All columns sum to 100 percent, and occupations categorized as "High Exposure" have values exceeding the median of their respective exposure index.

AI exposure and wages in the workforce. We next quantify wage gradients. Figure 2 relates average wages and exposure across occupational groups and shows a clear positive association: groups such as Managers, Professionals, and Technicians combine higher exposure with higher wages, while groups such as Elementary occupations and several manual categories combine lower exposure with lower wages.

Figure 2: Relation between AI Exposure and wages by occupational groups



Notes: Source: Authors' calculations based on GEIH (2023). This figure illustrates the AI exposure measure across 9 occupation groups. Marker sizes correspond to the proportion of workers for each group.

We complement this descriptive evidence with conditional correlations. We estimate log-wage regressions of the form

$$\log(w_i) = \alpha + \theta \mathbb{1}\{\text{HighExposure}_{o(i)}\} + X_i'\beta + \varepsilon_i, \tag{1}$$

where w_i denotes the wage of worker i, $\mathbb{1}\{\text{HighExposure}_{o(i)}\}$ is an indicator for whether worker i is employed in a high-exposure occupation $o(i)$, and X_i includes age, gender, education, informality, employment status, hours, and fixed effects for sector and department. In alternative specifications, we replace the indicator with the continuous index and its quadratic term to allow for non-linearities.

Table 5 shows that high-exposure occupations are associated with higher wages across indices, although magnitudes differ. These differences align with index construction. Complementarity-adjusted exposure yields smaller conditional wage gaps relative to raw applicability measures, consistent with the notion that the complementarity adjustment dampens exposure in the highest human-capital occupations. AI-Col displays a relatively large association when used at the occupational level, which is consistent with it capturing employer-requested capability content that is tightly linked to high-paying segments of the labor market. The results should be interpreted as correlations, not causal effects, but they establish that exposure measures map into economically meaningful wage gradients in Colombia.

Table 5: Wage differentials by AI exposure

	Webb (2020)			**Felten *et al.* (2021)**			**Pizzinelli *et al.* (2023)**			**AI-Col**		
	(1) Index	(2) Index + Index^2	(3) High exposure	(4) Index	(5) Index + Index^2	(6) High exposure	(7) Index	(8) Index + Index^2	(9) High exposure	(10) Index	(11) Index + Index^2	(12) High exposure
AI Index	0.003**	-0.004		0.131***	0.137***		0.127**	3.176***		0.239*	0.598***	
	(0.001)	(0.003)		(0.021)	(0.025)		(0.056)	(0.893)		(0.131)	(0.138)	
AI Index^2		0.000**			0.021			-0.282***			-0.073***	
		(0.000)			(0.024)			(0.083)			(0.019)	
High exposure			0.111*			0.172***			0.081*			0.218***
			(0.058)			(0.048)			(0.043)			(0.043)
Constant	12.694***	12.805***	12.769***	12.961***	12.929***	12.777***	12.148***	3.960	12.775***	12.799***	12.801***	12.822***
	(0.224)	(0.218)	(0.213)	(0.187)	(0.199)	(0.205)	(0.352)	(2.369)	(0.209)	(0.209)	(0.203)	(0.200)
Department FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	343,175	343,175	343,175	343,175	343,175	343,175	343,175	343,175	343,175	343,175	343,175	343,175
R-squared	0.463	0.464	0.462	0.469	0.469	0.464	0.461	0.463	0.460	0.463	0.469	0.469

Notes: Source: Authors' calculations based GEIH 2023. Clustered standard errors by occupation group in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Occupations categorized as "High Exposure" have values exceeding the median of their respective exposure index. In addition to including sector and department fixed effects, the estimates include variables such as age, gender, education, informality status, occupational group and hours worked.

Decomposing wage gaps between high- and low-exposure jobs. To understand what accounts for the wage differential between exposure groups, we implement Oaxaca–Blinder decompositions following the standard two-group formulation. The

decomposition separates the mean log-wage gap into a component explained by observable characteristics and a residual component associated with differences in coefficients (wage premia) across groups:

$$\overline{\log w}_H - \overline{\log w}_L = \left(\overline{X}_H - \overline{X}_L\right)\hat{\beta}_H + \overline{X}_L\left(\hat{\beta}_H - \hat{\beta}_L\right). \tag{2}$$

A robust pattern is that the unexplained component accounts for a large share of the overall gap. Observable characteristics, especially education and formal employment, explain an economically meaningful fraction of the wage differential, but substantial residual premia remain. This finding is consistent with exposure capturing bundles of capabilities and job attributes that are only partially observed in standard labor force surveys, and it is consistent with segmentation in which high-exposure jobs pay systematically higher premia even after conditioning on observables.

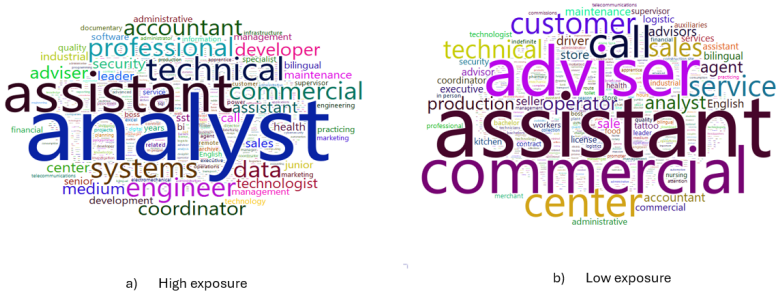
Where is AI-related capability demand concentrated in vacancies? We now return to the labor-demand lens. Table **??** compares the distribution of vacancies classified as high versus low exposure across indices. The contrast is stark: while occupation-based indices classify a substantial share of vacancies as high exposure, AI-Col classifies only a small fraction as exposed because it requires at least one AI-related term in the vacancy text. This divergence follows directly from measurement. Occupation-based indices assign the same score to all postings mapped into a given occupation, while AI-Col allows within-occupation variation and picks up only those postings that explicitly request AI-related tools, techniques, or closely related capabilities.

Table **??** shows that, under all indices, Professionals account for the largest share of high-exposure postings, and under AI-Col this concentration is particularly pronounced. The remaining distribution differs by index in informative ways. For example, clerical, service, and elementary categories are frequently classified as higher exposure under occupation-based indices than under AI-Col. This pattern is consistent with a setting in which AI-related capability demand is concentrated in digitally intensive, high-skill hiring, while broader occupational content may be potentially affected through indirect adoption channels that are not necessarily articulated in vacancy texts.

Text evidence on capability bundles: word clouds and term families. To characterize the skill content associated with exposure categories, we summarize vacancy text using two complementary approaches. First, Figure 3 contrasts word clouds of job titles for high- and low-exposure postings. High-exposure titles emphasize digital, managerial, and specialized technical content, while low-exposure titles emphasize operational roles, customer-facing services, and routine assistance positions. Second, we group frequent terms into skill families using the ESCO taxonomy to provide a structured interpretation.

The ESCO mapping highlights that high-exposure postings disproportionately mention ICT tools and data-related competencies, while low-exposure postings emphasize manual and operational tasks and basic service functions.

Figure 3: Word Clouds of Job Titles by Level of AI Exposure



Notes: Source: Authors' calculations based on job postings database (2022, 2023).

Table 6 complements this evidence by reporting mean AI-Col among exposed vacancies by 1-digit ISCO group and listing the most salient terms in each group. Professionals and Managers exhibit the highest mean intensity and are characterized by terms linked to data infrastructure and analytics (for example database tools, cloud environments, and programming languages). Technicians show intermediate intensity and comparatively more emphasis on digitalization and automation language. Lower-skill groups show lower intensity and fewer AI-specific terms, with occasional mentions that are consistent with partial adoption of digital tools for information handling and operational support.

Table 6: Composition of High AI Exposure Vacancies by Occupation Group According to AI-Col

Occupation (ISCO-08 major group)	**Mean AI-Col**	**Relevant terms (frequency)**
Managers	0.35	Power BI (409), IT (361), SQL (338), Software (279), Statistics (228), Cloud (171), Linux (152), Digitalization (141), SQL Server (134), Forecasting (116)
Professionals	0.65	SQL (3207), Power BI (2093), IT (1379), Software (1156), SQL Server (1147), Statistics (1117), Java (1043), Python (1001), Cloud (957), Linux (836)
Technicians	0.20	IT (915), Software (581), Digitalization (533), Power BI (492), SQL (489), Linux (478), Automation (468), Statistics (372), Cloud (210), Forecasting (136)
Clerical Support	0.10	IT (557), Digitalization (445), Software (215), Power BI (191), Statistics (126), SQL (125), Cloud (90), Forecasting (67), Automation (48), Statistical Analysis (32)
Services and Sales	0.06	IT (584), Digitalization (148), Software (81), Cloud (63), Statistics (58), Cybersecurity (49), Forecasting (33), QlikView (30), Linux (28), Digital Assistance (26)
Skilled Agricultural	0.05	Cloud (6), Statistics (5), IT (5), Digitalization (3), Machine Learning (3), Microsoft Azure (3), Pipeline (3), Power BI (3), Quantitative Analysis (2), Big Data (1)
Craft and Related Trades	0.05	Automation (68), Statistics (35), IT (31), Digitalization (23), Software (21), Power BI (18), SQL (12), Google Analytics (9), Java (6), SQL Server (6)
Machine Operators	0.02	IT (48), Automation (16), Software (15), Digitalization (7), Pandas (5), Linux (3), Cloud (3), Forecasting (3), SQL (2), Autonomous Vehicle (2)
Elementary Occupations	0.03	IT (23), Statistics (6), Power BI (6), SQL (6), Software (4), Statistical Analysis (2), Data Center (2), Digitalization (2), Pipeline (2), Automation (1)
Average AI-Col index	**0.17**	

Notes: Source: Authors' calculations based on job postings database (2022, 2023). The frequency of each term in each occupational group is shown in parentheses.

Offered wages in vacancies and exposure measures. Finally, we relate exposure to posted wages in vacancy data. Table 7 reports regression estimates analogous to equation (1),

replacing worker controls with vacancy controls (education requirements and 1-digit ISCO fixed effects). The occupation-based indices proposed by Felten and Pizzinelli display positive and statistically significant associations with offered wages. Webb's measure yields smaller and less stable estimates, consistent with it capturing dimensions related to automation potential rather than capability premia. AI-Col is also positively associated with offered wages, and specifications with a quadratic term suggest concavity, consistent with diminishing marginal wage differences at very high measured intensity.

Table 7: Wage differentials by AI exposure

	Webb (2020)			**Felten *et al.* (2021)**			**Pizzinelli *et al.* (2023)**			**AI-Col**		
	(1) Index	(2) Index + Index^2	(3) High exposure	(4) Index	(5) Index + Index^2	(6) High exposure	(7) Index	(8) Index + Index^2	(9) High exposure	(10) Index	(11) Index + Index^2	(12) High exposure
AI Index	0.0001***	-0.0000		0.02***	0.03***		0.03***	0.03**		0.01***	0.01***	
	(0.0000)	(0.0000)		(0.0004)	(0.0005)		(0.001)	(0.01)		(0.0001)	(0.0001)	
AI Index^2		0.0000***			-0.01***			-0.0003			-0.0001***	
		(0.0000)			(0.0004)			(0.001)			(0.0000)	
High exposure			-0.001***			0.02***			0.02***			0.11***
			(0.0004)			(0.0005)			(0.0004)			(0.001)
Education requested	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ISCO-08 (1 digit)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,169,943	1,169,943	1,169,943	1,169,943	1,169,943	1,169,943	1,169,943	1,169,943	1,169,943	1,169,943	1,169,943	1,169,943
R-squared	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16

Notes: Source: Authors' calculations based on job postings database (2022, 2023). Clustered standard errors by occupation group in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Taken together, the integrated evidence supports three conclusions. First, AI-related capability demand in posted vacancies is currently limited in prevalence but highly concentrated in high-skill occupational groups. Second, potential exposure indices imply broader technological applicability than what is currently articulated in vacancy texts, highlighting the importance of distinguishing applicability from revealed adoption signals in high-informality labor markets. Third, both vacancy-based and worker-based evidence point to meaningful wage gradients, and decomposition results indicate that observable characteristics explain part of the wage gap, but sizable residual premia remain. These patterns motivate the use of a country-specific, vacancy-based measure such as AI-Col as a complement to internationally used occupation-based indices.

4 Conclusion

This paper characterizes AI-related capability demand and occupational exposure in Colombia by combining a large corpus of online job postings with nationally representative household microdata. We begin from labor demand. Using 1,169,943 vacancies posted between March 2022 and December 2023, we construct two hiring-margin indicators: an

explicit AI-mention measure based on core terms and a continuous vacancy-level index, AI-Col, derived from a curated dictionary applied to vacancy titles and descriptions with inverse-document-frequency weighting. The first demand-side fact is that explicit AI-related requirements remain uncommon in posted hiring. At the same time, among the subset of exposed vacancies, AI-Col is highly right-skewed, indicating substantial heterogeneity in the intensity of AI-related capability requirements across postings.

A second demand-side result concerns measurement. When internationally used occupation-based exposure indices are assigned to the same vacancy data through harmonized occupational crosswalks, they imply substantially broader potential exposure than what is revealed in vacancy texts. This divergence is not an inconsistency; it reflects that the two families of measures capture different objects. Occupation-based indices quantify technological applicability embedded in occupational task or ability content, whereas vacancy-based measures capture what employers articulate when recruiting. In a high-informality labor market, where online postings disproportionately reflect formal and higher-skill recruitment channels, the vacancy-based measures are most informative for within-sample comparisons across occupations and requirements, rather than for interpreting economy-wide adoption levels.

We then connect these hiring-margin signals to the workforce using the 2023 wave of the GEIH, representing 22,520,713 employed workers. Across occupation-based exposure indices, a substantial share of workers is classified as highly exposed under a median-split definition, although levels vary by index. These differences are economically meaningful because the indices encode distinct conceptual emphases, ranging from automation-oriented applicability to complementarity-adjusted exposure. Worker characteristics display systematic sorting: high-exposure occupations are more likely to be formal and salaried and are staffed by more educated workers. When AI-Col is aggregated to occupations and mapped to workers, the resulting ranking emphasizes professional and technical groups and provides a locally grounded complement to international benchmarks.

Wage patterns are consistent with this demand-first characterization. In the vacancy data, exposure measures are positively associated with posted wages, and the premium is most pronounced when exposure is defined using vacancy text. In household microdata, workers in higher-exposure occupations earn more on average, and conditional regressions preserve positive associations between exposure and wages. Oaxaca–Blinder decompositions indicate that observable characteristics such as education and formality explain part of the wage gap between high- and low-exposure jobs, but a substantial residual component remains. We interpret these gradients as descriptive evidence that AI-related capability demand and exposure are concentrated in higher-paying segments

and are associated with wage premia that are not fully accounted for by standard observables, rather than as causal estimates of AI returns.

The combined evidence yields a clear implication for empirical work on developing economies. Occupation-based exposure indices remain useful to benchmark potential applicability, but they should not be interpreted as direct measures of adoption, particularly in contexts with uneven diffusion. Vacancy-based measures capture a distinct, policy-relevant margin: revealed capability demand at the point of hiring. Using both perspectives jointly is therefore essential to distinguish where AI could matter from where it is already being priced and signaled in labor demand.

Several limitations qualify the interpretation. Online postings do not represent the full hiring process, especially in informal segments; occupational coding from text and crosswalk harmonization introduce measurement noise; and the analysis is descriptive rather than causal. These limitations motivate next steps that build on the same framework: tracking vacancy signals over time to study diffusion, linking vacancy-based measures to firm characteristics where feasible, and combining hiring-margin evidence with direct adoption measures from firm surveys or administrative sources. Notwithstanding these constraints, the results show that integrating vacancy text with household microdata provides a disciplined basis to study AI-related labor market change in a high-informality setting, and that the gap between potential exposure and revealed capability demand is itself an informative object.

References

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1):S293–S340.

Acemoglu, D. and Restrepo, P. (2019). How Technology Displaces and Reinstates Labor. (12293):1–35.

Alekseeva, L., Azar, J., Giné, M., Samila, S., and Taska, B. (2021). The demand for AI skills in the labor market. *Labour Economics*, 71(October 2019):102002.

Alonso, C., Berg, A., Kothari, S., Papageorgiou, C., and Rehman, S. (2022). Will the AI revolution cause a great divergence? *Journal of Monetary Economics*, 127:18–37.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the u.s. labor market. *American Economic Review*, 103(5):1553–1597.

Babashahi, L., Barbosa, C. E., Lima, Y., Lyra, A., Salazar, H., Argôlo, M., de Almeida,

M. A., and de Souza, J. M. (2024). AI in the Workplace: A Systematic Review of Skill Transformation in the Industry. *Administrative Sciences*, 14(6).

Baruffaldi, S., Van Beuzekom, B., Dernis, H., Harhoff, D., Rao, N., Rosenfeld, D., and Squicciarini, M. (2020). Identifying and measuring developments in artificial intelligence : Making the impossible possible. *OECD Science, Technology and Industry Working Papers*, pages 1–68.

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.

Calvino, F., Samek, L., Squicciarini, M., and Morris, C. (2022). Identifying and characterising ai adopters: A novel approach based on big data.

Carbonero, F., Davies, J., Ernst, E., Fossen, F. M., Samaan, D., and Sorgner, A. (2023). *The impact of artificial intelligence on labor markets in developing countries: a new method with an illustration for Lao PDR and urban Viet Nam*, volume 33.

Cazzaniga, M., Jaumotte, M., F., Li, L., Melina, M., and Panton, A. J., Pizzinelli, C. (2024). *Intelligence and the future of artificial intelligence*.

Copestake, A., Pople, A., Marczinek, M., and Stapleton, K. (2024). AI and Services-Led Growth : Evidence from Indian Job Adverts .

Dawson, N., Rizoiu, M.-A., and Johnston, D. B. (2021). Changing Labour Market Dynamics in Australia: Skill Shortages, Job Transitions, and Artificial Intelligence Technology Adoption. (June).

Dernis, H., Calvino, F., Moussiegt, L., Nawa, D., and Samek, L. (2023). Identifying artificial intelligence actors using online data. pages 1–50.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). Gpts are gpts: Labor market impact potential of llms. *Science*, 384(6702):1306–1308.

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.

Fontanelli, L., Calvino, F., Criscuolo, C., Nesta, L., and Verdolini, E. (2025). Human after all: Occupations at the core of ai adoption. *Labour Economics*, page 102754.

Frey, C. B. and Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological forecasting and social change*, 114:254–280.

Garcia-Suaza, A., Caiza, P., Sarango-Iturralde, A., and Varela, D. (2025). How green is labor demand? an analysis for colombia. Mimeo.

Gehlhaus, D. (2021). The U . S . AI Workforce Understanding the Supply of AI Talent. (January).

Georgieff, A. and Hyee, R. (2022). Artificial intelligence and employment: New cross-country evidence. *Frontiers in artificial intelligence*, 5:832736.

Gmyrek, P., Winkler, H., and Santiago, G. (2024). *Buffer or Bottleneck?Employment Exposure to Generative AI and the Digital Divide in Latin America*. Number July.

Green, A. (2024). Artificial Intelligence and the Changing Demand for Skills in the Labour Market. *Oecd Artificial Intelligence Papers*, 14(14).

Kouretsis, A., Bampouris, A., Morfiris, P., and Papageorgiou, K. (2020). *labourR: Classify Multilingual Labour Market Free-Text to Standardized Classifications*. R package version 1.0.0.

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

Mhlanga, D. (2021). Artificial intelligence in the industry 4.0, and its impact on poverty, innovation, infrastructure development, and the sustainable development goals: Lessons from emerging economies? *Sustainability (Switzerland)*, 13(11).

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192.

Pizzinelli, C., Panton, A., Tavares, M. M., Cazzaniga, M., and Li, L. (2023). Labor Market Exposure to Differences and Distributional Implications.

Prytkova, E., Petit, F., Li, D., Chaturvedi, S., and Ciarli, T. (2024). The employment impact of emerging digital technologies.

Sarango-Iturralde, Alexander, Gallego, J., and Montaña, M. (2021). Impacto del COVID-19 en la demanda de habilidades en ocupaciones teletrabajables en Colombia.

Squicciarini, M. and Nachtigall, H. (2021). Demand for AI skills in jobs.

Su, Z., Togay, G., and Côté, A. M. (2021). Artificial intelligence: a destructive and yet creative force in the skilled labour market. *Human Resource Development International*, 24(3):341–352.

Theofanos, M., Choong, Y.-Y., and Jensen, T. (2024). AI use taxonomy: A human-centered approach.

Webb, M. (2020). The impact of artificial intelligence on the labor market. *Available at SSRN 3482150*.

5 Appendix

A Calculation of international indices

A.1 Felten (2021) index:

The exposure index is defined as follows:

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

$$AIOE_k = \frac{\sum_{j=1}^{52} A_{ij} L_{jk} I_{jk}}{\sum_{j=1}^{52} L_{jk} I_{jk}}$$

In this formula, i represents an AI application (EFF), j corresponds to an occupational skill from the O*NET taxonomy, and k refers to the specific occupation. The term A_{ij} denotes the skill-level exposure score, which measures the extent to which skill j is impacted by AI application i. To calculate the exposure score for each occupation, A_{ij} is weighted by two occupation-specific factors: the prevalence L_{jk}, which indicates how commonly skill j is used in occupation k, and the importance I_{jk}, which reflects the criticality of skill j to that occupation. By multiplying the skill-level exposure with these weights and summing across all 52 skills, the index captures the cumulative AI exposure for an occupation. The denominator ensures proper scaling, resulting in a normalized measure that represents the average weighted exposure of each occupation to AI. This approach provides a comprehensive framework to link AI advancements to occupational skill structures, highlighting the nuanced impact of technology on the labor market.

A.2 Webb (2020) index:

This process assigns each pair to a higher-order hierarchical group, facilitating a more structured and interpretable analysis. To calculate the index, consider a given technology $t \in T$. Let f_c^t denotes the raw count of occurrences of the aggregated verb-noun pair c as extracted from the patent titles associated with. The set of all aggregated verb-noun pairs for technology t is denoted as C^t . The relative frequency rf_c^t of a specific verb-noun pair c within the patent titles for technology t is then calculated to reflect the prominence of that functional relationship. This methodology provides a systematic way to link technological innovations to their underlying functions and applications.

$$rf_c^t = \frac{f_c^t}{\sum_{c \in C_t} f_c^t} \tag{3}$$

$$\text{Exposure}_{i,t} = \frac{\sum_{k \in K_t} \left[w_{k,i} \cdot \sum_{c \in S_k} rf_c^t \right]}{\sum_{k \in K_t} \left[w_{k,t} \cdot \{c : c \in S_k\} \right]}$$

Thus, in the Exposure index, K_i is the set of tasks in occupation i , and S_k is the set of verb-noun pairs extracted from task $k \in K$. Finally, $w_{k,i}$, the weight of task k within occupation i , is an average of the frequency, importance, and relevance of task k for occupation i , as specified in the O*NET database, with weights scaled to sum to 1.

A.3 Pizzinelli (2023)

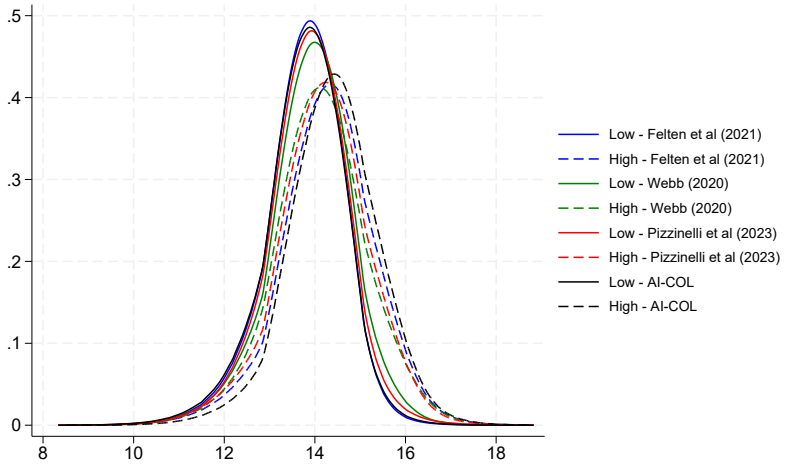
The complementarity component incorporates physical and social factors that assess the possibility of severe consequences from errors, alongside the level of education and training required to perform a given occupation. The resulting index, called C-AIOE is defined as follows:

$$\text{C-AIOE} = \text{AIOE}_k \left(1 - (\theta_i - \theta_{\min})\right)$$

Where θ is the complementarity index. This means that as the level of complementarity increases, the penalty on exposure also increases.

B Figures

Figure B.1: Wage distribution by AI Exposure



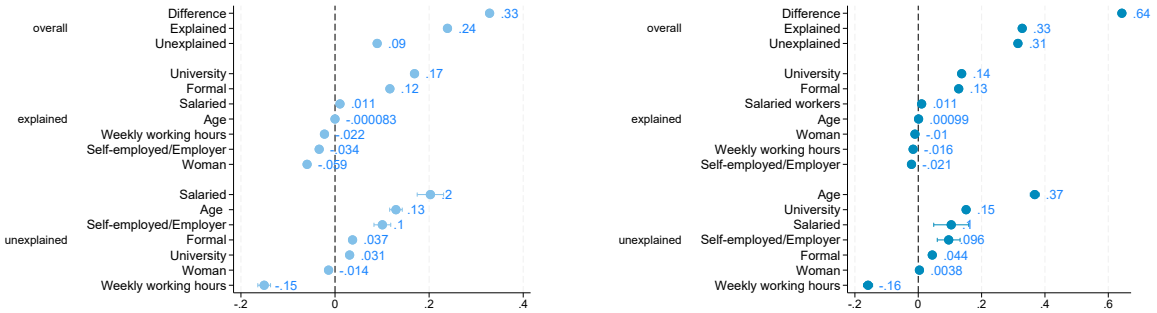
Notes: The occupations categorized as "High Exposure" have values exceeding the median of their respective exposure index. All calculations were performed using expansion factors from the 2018 census. Source: Authors' calculations based on GEIH (2023).

Figure B.2: Task shares by formality and sector



(a) Felten et al. (2021)

(b) Webb (2020)



(c) Pizzinelli et al. (2023)

(d) AI-Col

Notes: Occupations categorized as "High Exposure" have values exceeding the median of their respective exposure index. The estimates include sector fixed effects. Source: Authors' calculations based GEIH 2023.

C Tables

Table C.1: AI-Col dictionary: term families and illustrative examples

Term family	**Illustrative terms (examples)**
Core AI / GenAI	AI; ChatGPT; OpenAI; prompt; RLHF; hallucination
ML/DL methods	machine learning; deep learning; XGBoost; SVM; PCA; k-means
Neural architectures	CNN; RNN; LSTM; BERT; GAN; autoencoder; VAE
NLP / text processing	NLP; word2vec; POS tagging; OCR; translation
Data & analytics stack	SQL; PostgreSQL; MongoDB; ETL; Tableau; SSIS/SSRS/SSAS
Compute / cloud / infra	GPU; TPU; AWS; EC2; S3; PaaS; DevOps; Linux
Automation & applied AI	RPA; chatbot; pipeline; hyperautomation
Governance / risk / metrics	bias/sesgo; discrimination; AUC/ROC; SMOTE; overfitting/underfitting

Notes: Terms are shown for illustration only; the full bilingual dictionary (Spanish/English) is provided in the supplementary materials. Terms were curated to minimize false positives from generic business or IT language.