The EPOCH of AI: Human-Machine Complementarities at Work

Isabella Loaiza[a] and Roberto Rigobón[a]

[a]MIT Sloan School of Management

October 1, 2025

Abstract

We introduce the EPOCH framework (Empathy, Presence, Opinion, Creativity, and Hope) to capture human capabilities that complement, rather than substitute, artificial intelligence. Using network-based methods that map task interdependencies across all US occupations, we develop three metrics: (i) an EPOCH score measuring human-intensive skills, (ii) a potential-for-augmentation score, and (iii) a risk-of-substitution score. This framework explicitly distinguishes AI's roles in augmenting versus automating work, addressing a key gap in the literature. Our results show a clear shift toward more human-intensive work. New tasks emerging in 2024 carry significantly higher EPOCH scores than pre-existing tasks, and high-EPOCH tasks are performed more frequently. At the occupational level, EPOCH-intensive jobs experienced stronger employment growth from 2015 to 2023, higher hiring rates in 2024, and more favorable projections through 2034. In contrast, occupations with higher substitution risks show consistently negative outcomes across past employment, current hiring, and future projections. Finally, augmentation scores are negatively associated with recent employment and hiring trends, but show no significant link to long-run employment projections.

Augmentation | Automation | Future of Work | Economic Networks

1 Introduction

echnological progress often advances incrementally, but some inventions trigger discontinuous leaps that reshape history. Money, the wheel, writing, engines, electricity, computers, and the internet exemplify such General-Purpose Technologies (GPTs) [11, 32, 39]. A GPT is defined by three features: it improves over time, spurs complementary innovations, and transforms multiple economic and social domains through widespread use [10]. Early versions typically differ greatly from their mature forms, evolving through sustained refinement and collective experimentation. These technologies create and reshape markets, displace older ones, and drive far-reaching societal change.

The transformative impact of GPTs is shaped by the political, social, economic, and institutional contexts in which they emerge. The same technology can thus have vastly different effects across cities [28], nations [8], industries [16], firms [24], and workers [22]. Industries, firms, and workers that complement the technology typically benefit through productivity gains and job creation, while those that compete with it face displacement and market exit.

Uncertainty over these distributional effects often fuels resistance and fear. From the Luddite movement to today's workforce concerns about Artificial Intelligence (AI), each new GPT provokes similar anxieties. Economists [48] and technologists [30] regard AI as the latest GPT, with a widely cited 2013 study estimating that 47% of US occupations were at high risk of automation [29]. This finding prompted numerous efforts to measure AI's potential for labor substitution across the United States and Europe [20, 14, 22, 50, 45, 5, 44, 26].

However, an overemphasis on AI's labor-saving potential risks obscuring our understanding of diverse potential futures for work. This substitution focus appears in both theory—exemplified by metrics measuring automation risk but not augmentation across occupational categories—and practice, through instances of 'so-so' automation in organizational processes [2]. Just as GPTs can displace workers, they can also unlock

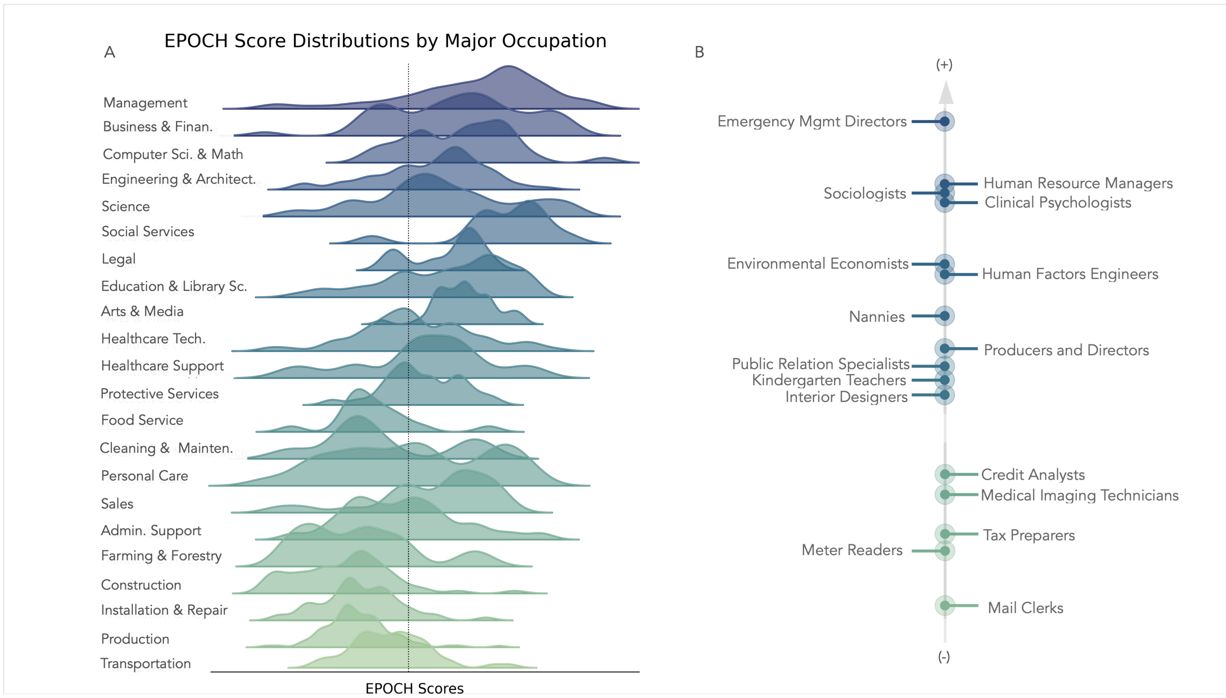

Figure 1: This figure shows the distribution and ranking of occupations by EPOCH Score. Panel A shows the EPOCH score distribution of all detailed occupations in O*NET's major occupational group. Panel B shows selected occupations ranked from the highest EPOCH score from top to bottom.

new human potential, turning teams of people and machines into "superminds" [41]. Yet most research has focused on automation risk rather than complementarity [49, 33].

This paper addresses this gap by developing three metrics to capture both AI's effects on labor and the distinctive strengths of human workers where AI falls short. Specifically, we ask: Which human capabilities complement AI's limitations? By applying these metrics to US occupational tasks and categories, we aim to promote a more balanced narrative about AI and work—one that places human capabilities at the center of the future economy.

The Frontiers of Automation

Despite significant advances in AI research, today's systems cannot match human performance across many real-world tasks [23]. Artificial General Intelligence remains elusive, and fundamental limitations may persist for years to come. We identify five core challenges that define the current frontiers of automation—areas where human capabilities remain essential.

Small Data Inference. Current AI relies on statistical learning from large datasets to identify patterns. This approach fails when data is insufficient, biased, rendering many observations equivalent to just a few, or when rare events must be predicted. Humans excel here by understanding causal mechanisms rather than relying solely on correlations. For instance, while machine learning cannot predict solar eclipses from observational data alone due to data scarcity, humans can because we grasp the underlying celestial mechanics [40].

On a practical level, the "small data" challenge persists because of the high costs of data collection, curation, and annotation. Surveys show that 96% of enterprises face data challenges, including quality and labeling. As a result, data scientists spend nearly twice as much time data cleaning as on model training, selection, and deployment [37].

Extrapolation Beyond Training Boundaries. Machine learning assumes that training data and future predictions share the same underlying distribution. State-of-the-art AI excels at interpolation but struggles when asked to extrapolate beyond its training boundaries, with performance degrading as the distance from training data increases [23]. This limitation affects both convergent thinking (finding single correct solutions, as in mathematical reasoning [47]) and divergent thinking (tackling open-ended problems requiring flexible knowledge application). In medicine, approximately 15% of patient records require extrapolation, often exceeding acceptable clinical thresholds [15].

Extrapolation challenges also affect AI's ability to engage in divergent thinking—tackling open-ended problems that require flexible application of knowledge across different contexts. In these cases, machines struggle to transfer skills learned in one domain to another, a limitation actively studied in transfer learning research. For instance, in natural language processing, models often have difficulty handling ambiguity. Similarly, in artistic and creative domains, AI's output tends to be closely tied to existing material, leading some to argue that its performance in generating content such as poems often lacks diversity and originality [37].

Multiple Justifiable Solutions. This challenge arises when problems allow for two or more valid but conflicting solutions, involve significant ambiguity or tacit knowledge, or are fundamentally indeterminate. From a technical standpoint, [9] argues that algorithms often generate a single solution, potentially obscuring equally viable alternatives and biasing available options. From the perspective of complex systems and embodied cognition, people and social systems are inherently indeterminate and unpredictable. Some scholars, therefore, contend that, given the indeterminate nature of human beings and social systems, AI's predictive abilities in social contexts may not meaningfully exceed those of human judgment [43].

Moral dilemmas, or "wicked problems," illustrate situations where multiple solutions can be justified. According to [19], "wickedness" stems from both ambiguity in the facts defining a problem and a lack of consensus on normative criteria for its resolution, making such problems difficult to encode in current AI systems. Beyond technical limitations, some argue that even if these challenges were overcome, ontological constraints would still limit AI's applicability in contexts like moral dilemmas [21]. Moreover, in some cases, the validity of a solution depends on the decision-making process itself, not just the outcome.

Relational Outcomes. Some decision-making processes aim not at specific answers but at building connections, relationships, or shared experiences among participants. AI's lack of Theory of Mind limits its ability to understand social cues, interpret implicit communication, or engage in genuinely bidirectional relationships [23]. Building authentic human connections requires capacities like empathy and compassion that extend beyond pattern recognition.

Subjective Beliefs and Value-Driven Decisions. The final challenge highlights that individuals often make decisions based on outcomes that differ from what the data suggests. While this may appear to be an error in judgment, it frequently is not. Some of the most transformative decisions in history—such as movements for civil rights and women's rights—were driven by beliefs that defied the status quo, even when prevailing data seemed to support it. Statistical learning and inference from data alone are therefore insufficient; it is also necessary to recognize the causal relationships behind the measured data.

Algorithmic fairness lies at the heart of the discrepancy between sample distribution and subjective perspectives. For example, when a group is systematically excluded from education, the data will inherently reflect this bias, offering no empirical evidence to predict the potential of those denied opportunities. Decisions that challenge both the status quo and biased data often stem from a conviction that the current situation is unjust. A similar issue arises when sample data suffers from omitted variable bias, yet the decision maker accounts for the unseen factors. In summary, maintaining a belief distribution that differs from the sample data is not inherently wrong; in some cases, it can provide a more just or insightful perspective.

To see how AI's limitations map to EPOCH capabilities see Table S3 in the SI.

The Foundations of Augmentation: Human Capabilities

The previous subsection examined AI's limitations in certain contexts. Humans face similar challenges but have developed structured procedures, shared responsibilities, and social norms to navigate them. We refer to this combination of individual skills and societal institutions as human capabilities.

Skills are a fundamental component of capabilities, but they are not equivalent. Skills enable individuals to perform specific tasks effectively but are narrowly defined. Capabilities encompass broader qualities that allow individuals to integrate and apply skills across diverse, open-ended contexts [31]. The capabilities in our framework are distinct from so-called "soft skills," a misleading term suggesting these abilities are easy to develop. In reality, they are often among the most challenging to cultivate and teach.

Based on expert interviews, we identified five groups of human capabilities that enable work in areas where machines are limited. These form the acronym EPOCH:

- Empathy and Emotional Intelligence
- Presence, Networking, and Connectedness
- Opinion, Judgment, and Ethics
- Creativity and Imagination
- Hope, Vision, and Leadership

Each EPOCH capability is not universally beneficial—history shows instances where these have been used detrimentally. The relevant question is not whether these capabilities are inherently good or bad, but whether they can be substituted and whether humans would view such substitutions as preferable.

When describing each group, we provide examples rather than formal definitions to illustrate where AI can assist versus where it falls short. For detailed descriptions of specific capabilities within each group refer to the section *EPOCH definitions* in the SI.

The Underlying Structure of Automation and Augmentation

The task-based framework developed by Autor and colleagues has been instrumental in understanding how technological change shapes work [7, 1]. This framework demonstrates that changes occur at the level of tasks—the fundamental units of work—rather than entire occupations. Automation reallocates tasks from workers (labor) to machines (capital), thereby shifting the content of work within occupations [2].

Table 1: EPOCH Score for New, Current, Retired Tasks and Task Frequency ChangesS (2016–2024)

	Current/Old	New/Current	New/Retired	Δ Frequency 2016–2024
Intercept	0.920***	0.038***	0.309***	0.029***
	(0.003)	(0.002)	(0.015)	(0.005)
EPOCH Score	0.026***	0.029***	0.256***	0.022*
	(0.007)	(0.006)	(0.038)	(0.013)
Observations	18,174	17,679	2,125	16,095
R^2	0.001	0.002	0.021	0.000
Adj. R^2	0.001	0.001	0.021	0.000

Note: *p<0.1; **p<0.05; ***p<0.01

Many studies on the future of work have focused on measuring *risk*: the risk of automation [29, 14, 28, 5, 45, 51, 25], the risk of unemployment [27], and the risk of increasing inequality [36]. Others have developed broader metrics to quantify "exposure to AI," encompassing both automation and augmentation [22, 50, 20, 46, 17]. However, to our knowledge, no study has attempted to analyze both effects separately from each other.

Augmentation, by contrast, is seen as a vital strategy for managing AI's impact on the labor force [13] and for restoring dignity, agency, and middle-class jobs [6]. Its measurement, however, has largely been limited to specific applications such as sales [38], call centers [12], and radiology [35]. Unsurprisingly, a recent meta-analysis [49] emphasized the need for a more comprehensive understanding of human-AI complementarities and highlighted tasks where human-machine synergies emerge experimentally. Our work builds on these insights by developing metrics that separately measure automation and augmentation for tasks in real-world settings.

Let's begin with the definition of augmentation. Labor augmentation occurs when using a machine in one task increases productivity in other tasks, enhancing overall labor productivity [3]. This definition underscores a key element for augmentation—*network effects*. Unlike automation, which transfers tasks directly from humans to machines, augmentation involves interactions among tasks, whether in pairs, clusters, or entire networks.

This perspective aligns with Autor, who notes that " tasks currently bundled into these jobs [middle-skill jobs] cannot be readily unbundled" [18]. " While the BLS categorizes tasks as 'Core' or 'Supplementary,' hinting at a task hierarchy, we argue that this binary classification fails to capture the nuanced structure of tasks within and across occupations.

The effects of automating specific tasks can vary dramatically depending on their location within the task network. Just as network structures influence the diffusion of innovations [42], shape occupational skill requirements [4], and determine national productive capacities [34], task networks play a critical role in their resilience to new technologies. Incorporating this network perspective allows for a better understanding and measurement of AI-driven augmentation. By adopting a network-based view of tasks, we develop a scalable metric for measuring augmentation, addressing a key gap in the future-of-work literature.

Results

Our analysis shows that as AI advances, the work performed by US workers is systematically shifting toward tasks that emphasize the human-intensive capabilities identified in the EPOCH framework. This pattern persists at both the task and occupational levels. It is visible in employment changes over the past decade, current hiring trends, and projected employment shifts in the coming decade. All the coefficients in our regressions are estimated with standardized data.

New tasks are more human-intensive than existing tasks, and retired tasks

To begin our task analysis, we categorized tasks from two O*NET versions. We used 2016 as the baseline and 2024 as the most recent update at the time of analysis. 'New' tasks are those appearing for the first time in 2024. 'Current' tasks are those present in both datasets, while 'Retired' tasks are those present in 2016 but removed in 2024. Our analysis shows that, compared to the 2016 O*NET version, 815 new tasks were added by 2024, 1310 tasks were retired, leaving 16,864 tasks classified as current. We then compared their EPOCH scores across these three categories.

Table **??** presents these findings based on three regression models comparing EPOCH scores across task groups. 'New' tasks entering the work pool score substantially higher than 'current' ($\beta = 0.029$, $p < 0.01$) and 'retired' tasks ($\beta = 0.256$, $p < 0.01$), while current tasks also have higher scores than retiring tasks ($\beta = 0.026$, $p < 0.01$). These results suggest that newer tasks have higher EPOCH scores, requiring more human-intensive capabilities to be performed. Given that EPOCH scores take values between 0 and 1, the coefficients, particularly the one for 'new' and 'retired' tasks, exhibit significant effects.

Human-intensive tasks have increased in frequency

O*NET reports the frequency with which workers perform each task on a seven-point categorical scale. Category 1 corresponds to tasks performed yearly or less, category 2 to tasks performed more than once a year, category 3 to tasks carried out more than once a month, category 4 to tasks performed more than weekly, category 5 to daily tasks, category 6 to tasks performed more than once daily, and category 7 to tasks performed hourly. Because the categories are monotonic with respect to frequency, we treat them as a discrete ordinal variable. To calculate the average frequency of each task, we computed the weighted average across the seven categories, using the number of incumbents reporting each category as weights.

Analysis of task frequency data shows that human-intensive tasks are generally performed less frequently than other workplace activities according to O*NET. However, these high EPOCH tasks are becoming more frequent in workers' routines between 2016 and 2024 as shown in Table 1 with the model titled Δ Frequency 2016-2024. Thus, these results suggest that workers are engaging in these human-intensive tasks more regularly.

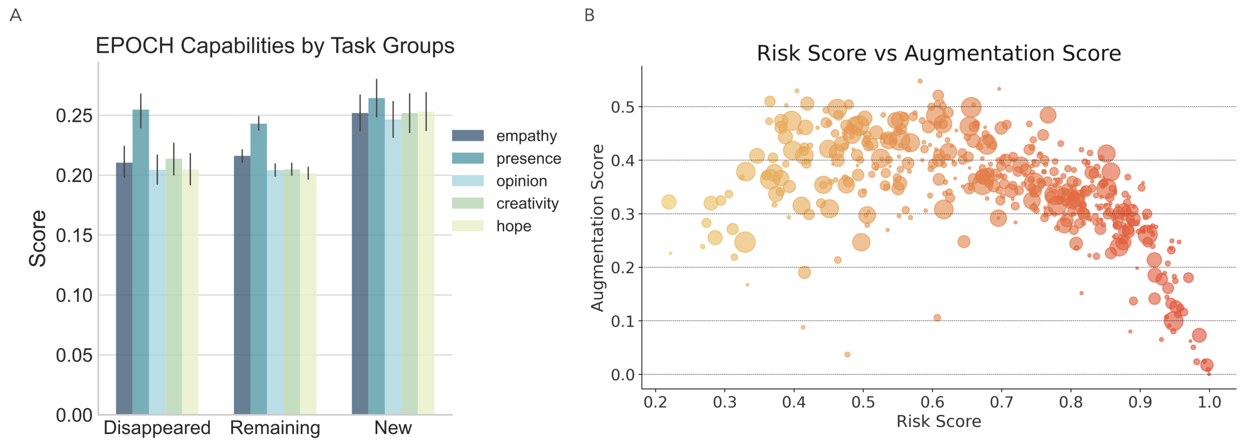

Figure 2: Panel A shows a bar chart shows individual EPOCH capability scores based on the three tasks groups New, Current and Retired. New tasks have higher scores for all five capabilities. Panel B shos the relationship between our Augmentation and Automation scores.

We acknowledge that this analysis has limitations due to O*NET's categorical frequency scale, which is censored at both ends with only seven categories. Nevertheless, even within this constrained measurement framework, we observe a meaningful shift toward more frequent performance of EPOCH-intensive tasks over a relatively short period.

Human-Intensive occupations experience more employment growth, hiring and better projections

These micro-level changes translate into an occupation-level shift towards more human work. We find that occupations with higher EPOCH scores have experienced significantly stronger job growth from 2015 to 2023 (the latest employment data at the time of analysis), a pattern that extends to 2025 hiring trends and future employment projections for the next decade.

First, from 2015 to 2023, occupations with higher EPOCH scores experienced substantially stronger employment growth, with a standard deviation in EPOCH score associated with 0.132 standard deviation increase in employment change. That is roughly equivalent to 12,000 additional jobs for a one standard deviation in EPOCH score. These results are shown in Table 2, where all coefficients were estimates using centered and standardized data. This pattern is consistent even if we regress individual EPOCH capability bins, with all the bins showing positive and statistically significant results as shown in Table S3.

Conversely, occupations with higher automation risk showed significant employment declines, consistent with AI substitution (β=-0.219, p<0.01), equating to almost 20,000 fewer jobs per EPOCH standard deviation. Our EPOCH framework demonstrates superior predictive power compared to existing AI exposure metrics, maintaining statistically significant associations with employment growth where other established measures show no relationship in our regressions (Table S4 in the SI). Occupations with higher augmentation scores also showed employment decline in the last decade (β=-0.136, p<0.01), albeit the effect was smaller than for occupations with high risk scores.

This historical pattern extends to current labor markets in 2025. Analysis of hiring data between January and July 2025 from Revelio Labs confirms that occupations with higher EPOCH scores have higher hiring rates than other occupational groups. It is worth noting, however, that these results are derived from hiring numbers at the 2-digit level SOC-codes, otherwise known as 'major occupation groups'. All other results are obtained with more granular, 6-digit SOC Codes.

Looking to the future, we find evidence that our scores are statistically associated with the employment projections from the BLS until 2034. As Table **??** shows, we find that occupations with higher EPOCH scores will continue to experience job growth in the upcoming decade (β=0.114, p<0.01), adding roughly 3,762 more jobs per EPOCH score deviation. In contrast, occupations with high risk scores are expected

Table 2: Relationship between Occupational EPOCH, Augmentation and Automation Scores with Past, Current and Future Employment

	Past Employment 15–23		Current Hires 2025		Future Projections 24–34	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.000	0.000	0.058	0.058	-0.000	-0.000
	(0.040)	(0.040)	(0.132)	(0.116)	(0.036)	(0.036)
EPOCH Score	0.132***		0.796***		0.114***	
	(0.040)		(0.050)		(0.036)	
Automation Score		-0.219***		-1.066***		-0.155***
		(0.047)		(0.063)		(0.043)
Augmentation Score		-0.136***		-0.357***		-0.071
		(0.047)		(0.063)		(0.043)
Observations	614	608	154	154	749	749
R^2	0.017	0.036	0.636	0.722	0.013	0.017
Adj. R^2	0.016	0.033	0.618	0.706	0.012	0.014

Note: *p<0.1; **p<0.05; ***p<0.01

to decline in employment (β=-0.155, p<0.01). Interestingly, we find no statistically significant relationship between occupations with high augmentation scores and employment projections.

The consistency across past performance, current hiring, and future projections suggests that the economy's reorganization around human capabilities represents a durable structural shift rather than a temporary adjustment to technological change.

Discussion

The contributions of this paper are threefold: (1) it measures augmentation alongside automation across all US occupations and introduces a new metric to evaluate human-intensive tasks; (2) it incorporates task interdependence as a critical factor in assessing AI's labor force impact using task co-occurrence networks, and (3) it leverages foundation models to calculate capability scores without requiring human labelers, thus reducing inter-rater variability, reporting bias, and evaluator subjectivity while enhancing scalability and cost-effectiveness.

Our analysis highlights a nuanced relationship between AI's potential impact on employment within the American labor force. The findings indicate that an increase in the EPOCH score, which signifies greater capability for workers to thrive alongside AI, correlates with a measurable increase in employment levels. Our augmentation scores indicate that while AI enhances productivity, it does not necessarily result in higher employment.

This research provides critical insights into AI's nuanced role in the labor market, advocating for integration that emphasizes human-AI complementarities to support a sustainable, equitable future of work. Finally, more research is needed to understand the impact of AI on wages, and how AI will impact different groups of workers in the country.

2 Methods

We develop three measures to quantify the potential impact of AI and emerging technologies on the workforce: 1) the EPOCH score capturing human-intensive capabilities, 2) an automation risk score, and 3) a potential for augmentation score. EPOCH scores are calculated at task, task-cluster, and occupation levels, while automation and augmentation scores are computed at the occupation level.

We use O*NET occupational data (2016 and 2023 releases) containing 19,000 unique tasks across 947 occupations, together with employment data from the BLS and hiring data for 2025 from Revelio Labs, all of which are publicly available sources.

Task clusters and the task-interdependence network

Following the task-based framework, we begin with task-level analysis. We clean task statements using a disambiguation procedure (**??** details in SI) that addresses duplicate task descriptions and minor semantic modifications across versions. Applying this procedure allows us to distinguish genuine changes in work from superficial semantic modifications, improving our ability to capture meaningful shifts in tasks and occupations. To reduce sensitivity to minor variations, we paraphrase each task statement 20 times using GPT-4.

We then estimate the task interdependence network TC_k for all tasks in the dataset. Since O*NET task statements are occupation-specific with limited overlap, we embed the 20 paraphrased versions of each task, get the average embedding, and aggregate these average embeddings into 750 task-clusters TC_k based on their cosine similarity. Robustness checks with 700-1,000 clusters yield similar results, and a table showing examples of task clusters is included in the SI. Then, to build the co-occurrence network, we compute the pairwise co-occurrence of each task-cluster across all occupations in O*NET. Occupation-specific networks are considered subgraphs of the overall task-interdependence network, where the subgraph contains only the subset of task-clusters and their connections present in that occupation.

The intuition behind using task co-occurrence to quantify the interdependence or complementarity between tasks is that if two task clusters, denoted by k_1 and k_2, are highly interdependent, they will consistently appear together. If k_1 and k_2 are perfect complements, whenever an occupation includes task cluster k_1, it also includes task cluster k_2.

EPOCH Score for Tasks Clusters

The EPOCH score for task-clusters uses a CES (Constant Elasticity of Substitution) function. The intuition is that having at least one EPOCH capability provides protection against automation, but achieving high scores in multiple dimensions is preferable to relying on just one. This multidimensional approach recognizes that while individual capabilities may provide resilience, a combination of strengths enhances adaptability and innovation, creating a workforce equipped to thrive alongside technological advancements.

To compute the EPOCH score for each the task-cluster, we first compute task-level scores for each EPOCH capability by measuring the cosine similarity between each task cluster embedding each capability embedding. Therefore, every task has a corresponding set of capability scores $\{e_j, p_j, o_j, c_j, h_j\}$. To aggregate individual capability scores into task-cluster EPOCH scores $EPOCH_{TC_k}$ for each cluster, we use the CES function shown in Equation 1.

$$EPOCH_{TC_k} = (w_e e^{\gamma} + w_p p^{\gamma} + w_o o^{\gamma} + w_c c^{\gamma} + w_h h^{\gamma})^{1/\gamma} \tag{1}$$

where

$$w_e + w_p + w_o + w_c + w_h = 1 \; and \; \gamma > 1 \tag{2}$$

Occupational EPOCH Scores

We compute occupational EPOCH scores $EPOCH_{O_i}$ by aggregating the EPOCH scores of individual tasks within each occupation. For each cluster-of-tasks we compute an EPOCH score with Equation 3.

$$EPOCH_{O_i} = \sum_{k \in K_{i,k}} w_{i,k} \cdot EPOCH_{CoT_k} \tag{3}$$

where $\sum w_k = 1$ and each $w_{i,k}$ is proportional to the number of tasks in the occupation that belong to the same cluster. In other words, if an occupation has 20 tasks, and 15 of them each one belongs to a single cluster, but the remaining 5 belong to a single cluster the last cluster has a weight five times larger than the other 15.

Occupational Risk and Augmentation Scores

In the computation of risk and augmentation we use task-cluster EPOCH scores and the subgraph of the task interdependence network that only contains the tasks belonging to that occupation. We denote the co-occurrence between task-cluster i and task-cluster j as C_{ij}.

In computing augmentation, we reward the complementarity of the tasks and the fact that one task may be substituted by AI while the other is not. Accordingly, the measure emphasizes differences in the EPOCH scores across pairs of co-occurring task clusters. We use to compute augmentation for each occupation n using Equation 5. For simplicity in the notation, since we focus only on task clusters, we omit TC_k and refer to each pair of task clusters simply as i or j.

$$p_{ij} = \frac{C_{ij}}{\sum_{i \neq j} C_{ij}} \tag{4}$$

$$Aug_n = \left(\sum_{i \neq j} p_{ij} \, EPOCH_{TC_i} - EPOCH_{TC_j}^{\gamma_{aug}} \right)^{1/\gamma_{aug}} \tag{5}$$

The risk measure for a task-clusters is directly one minus the EPOCH score. We also incorporate the co-occurrence task network to capture the idea that is that the risk increases when the tasks are disconnected. We compute risk using Equation 8. Just like in the previous equation, we use i and j to reference a pair of co-occurring task-clusters.

$$w_{r_i} = \frac{1}{K} \sum_{j=1}^{K} \left(1 - C_{ij}\right) \tag{6}$$

$$p_{r_i} = \frac{w_{r_i}}{\sum_i w_{r_i}} \tag{7}$$

$$Risk_n = \left(\sum_i p_{r_i} \left(1 - EPOCH_{TC_i}\right)^{\gamma_{risk}} \right)^{1/\gamma_{risk}} \tag{8}$$

where k is the number of task-clusters in the occupation.

References

[1] Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, volume 4, pages 1043–1171. Elsevier, 2011.

[2] Daron Acemoglu and Pascual Restrepo. Automation and new tasks: How technology displaces and reinstates labor. *Journal of economic perspectives*, 33(2):3–30, 2019.

[3] Ajay Agrawal, Joshua S Gans, and Avi Goldfarb. Do we want less automation? *Science*, 381(6654):155–158, 2023.

[4] Ahmad Alabdulkareem, Morgan R Frank, Lijun Sun, Bedoor AlShebli, César Hidalgo, and Iyad Rahwan. Unpacking the polarization of workplace skills. *Science advances*, 4(7):eaao6030, 2018.

[5] M Arntz. The risk of automation for jobs in oecd countries: A comparative analysis. *ECD Social, Employment and Migration Working Papers, No. 189*, 2016.

[6] David Autor. How ai could help rebuild the middle class, 2024. Accessed: 2024-11-08.

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

[8] Thereza Balliester, Adam Elsheikhi, et al. The future of work: a literature review. *ILO research department working paper*, 29:1–54, 2018.

[9] Abeba Birhane. Automating ambiguity: Challenges and pitfalls of artificial intelligence. *arXiv preprint arXiv:2206.04179*, 2021.

[10] Timothy F Bresnahan, Scott Stern, and Manuel Trajtenberg. Market segmentation and the sources of rents from innovation: Personal computers in the late 1980's, 1996.

[11] Timothy F Bresnahan and Manuel Trajtenberg. General purpose technologies: 'engines of growth.'. *Journal of Econometrics*, 65:83–108, 1995.

[12] Erik Brynjolfsson, Danielle Li, and Lindsey R Raymond. Generative ai at work. Technical report, National Bureau of Economic Research, 2023.

[13] Erik Brynjolfsson and Andrew McAfee. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies.* WW Norton & Company, 2014.

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

[15] Xuenan Cao and Roozbeh Yousefzadeh. Extrapolation and ai transparency: Why machine learning models should reveal when they make decisions beyond their training. *Big Data & Society*, 10(1):20539517231169731, 2023.

[16] Fulvio Castellacci. Structural change and the growth of industrial sectors: Empirical test of a gpt model. *Review of income and wealth*, 56(3):449–482, 2010.

[17] Emilio Colombo, Fabio Mercorio, Mario Mezzanzanica, and Antonio Serino. Towards the terminator economy: Assessing job exposure to ai through llms. *arXiv preprint arXiv:2407.19204*, 2024.

[18] H David. The "task approach" to labor markets: an overview. *Journal for Labour Market Research*, 46(3):185–199, 2013.

[19] Hans de Bruijn, Martijn Warnier, and Marijn Janssen. The perils and pitfalls of explainable ai: Strategies for explaining algorithmic decision-making. *Government information quarterly*, 39(2):101666, 2022.

[20] Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. Gpts are gpts: Labor market impact potential of llms. *Science*, 384(6702):1306–1308, 2024.

[21] Hubert Etienne. Solving moral dilemmas with ai: how it helps us address the social implications of the covid-19 crisis and enhance human responsibility to tackle meta-dilemmas. *Law, Innovation and Technology*, 14(2):305–324, 2022.

[22] Edward W Felten, Manav Raj, and Robert Seamans. Occupational heterogeneity in exposure to generative ai. *Available at SSRN 4414065*, 2023.

[23] Tao Feng, Chuanyang Jin, Jingyu Liu, Kunlun Zhu, Haoqin Tu, Zirui Cheng, Guanyu Lin, and Jiaxuan You. How far are we from agi. *arXiv preprint arXiv:2405.10313*, 2024.

[24] Chris Forman. The corporate digital divide: Determinants of internet adoption. *Management Science*, 51(4):641–654, 2005.

[25] World Economic Forum. The reskilling revolution: A future of jobs for all. Technical report, World Economic Forum, 2028. Accessed: 2024-11-08.

[26] Neil Foster-McGregor, Önder Nomaler, and Bart Verspagen. Job automation risk, economic structure and trade: a european perspective. *Research Policy*, 50(7):104269, 2021.

[27] Morgan Frank, Yong-Yeol Ahn, and Esteban Moro. Ai exposure predicts unemployment risk. *arXiv preprint arXiv:2308.02624*, 2023.

[28] Morgan R Frank, Lijun Sun, Manuel Cebrian, Hyejin Youn, and Iyad Rahwan. Small cities face greater impact from automation. *Journal of the Royal Society Interface*, 15(139):20170946, 2018.

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

[30] Avi Goldfarb, Bledi Taska, and Florenta Teodoridis. Could machine learning be a general purpose technology? a comparison of emerging technologies using data from online job postings. *Research Policy*, 52(1):104653, 2023.

[31] James J Heckman and Chase O Corbin. Capabilities and skills. *Journal of human development and capabilities*, 17(3):342–359, 2016.

[32] Elhanan Helpman. *General purpose technologies and economic growth*. MIT press, 1998.

[33] Jennifer L Heyman, Steven R Rick, Gianni Giacomelli, Haoran Wen, Robert Laubacher, Nancy Taubenslag, Max Knicker, Younes Jeddi, Pranav Ragupathy, Jared Curhan, et al. Supermind ideator: How scaffolding human-ai collaboration can increase creativity. In *Proceedings of the ACM Collective Intelligence Conference*, pages 18–28, 2024.

[34] César A Hidalgo, Bailey Klinger, A-L Barabási, and Ricardo Hausmann. The product space conditions the development of nations. *Science*, 317(5837):482–487, 2007.

[35] Brendan S Kelly, Conor Judge, Stephanie M Bollard, Simon M Clifford, Gerard M Healy, Awsam Aziz, Prateek Mathur, Shah Islam, Kristen W Yeom, Aonghus Lawlor, et al. Radiology artificial intelligence: a systematic review and evaluation of methods (raise). *European radiology*, 32(11):7998–8007, 2022.

[36] Anton Korinek and Joseph E Stiglitz. Artificial intelligence and its implications for income distribution and unemployment. In *The economics of artificial intelligence: An agenda*, pages 349–390. University of Chicago Press, 2018.

[37] Weixin Liang, Girmaw Abebe Tadesse, Daniel Ho, Li Fei-Fei, Matei Zaharia, Ce Zhang, and James Zou. Advances, challenges and opportunities in creating data for trustworthy ai. *Nature Machine Intelligence*, 4(8):669–677, 2022.

[38] Chengcheng Liao, Xin Wen, Shan Li, and Peiyuan Du. How effective is ai augmentation in human–ai collaboration? evidence from a field experiment. *Information Technology & People*, 2024.

[39] Richard G Lipsey, Kenneth I Carlaw, and Clifford T Bekar. *Economic transformations: general purpose technologies and long-term economic growth*. Oup Oxford, 2005.

[40] Klaus Mainzer and Reinhard Kahle. *Limits of AI-theoretical, practical, ethical*. Springer, 2024.

[41] Thomas W Malone. How human-computer 'superminds' are redefining the future of work. *MIT Sloan management review*, 59(4):34–41, 2018.

[42] Andrea Montanari and Amin Saberi. The spread of innovations in social networks. *Proceedings of the National Academy of Sciences*, 107(47):20196–20201, 2010.

[43] Arvind Narayanan. The promise and peril of artificial intelligence, 2019. Arthur Miller Lecture on Science and Ethics, Massachusetts Institute of Technology.

[44] Ljubica Nedelkoska and Glenda Quintini. Automation, skills use and training. *OECD Report*, 2018.

[45] Antonio Paolillo, Fabrizio Colella, Nicola Nosengo, Fabrizio Schiano, William Stewart, Davide Zambrano, Isabelle Chappuis, Rafael Lalive, and Dario Floreano. How to compete with robots by assessing job automation risks and resilient alternatives. *Science robotics*, 7(65):eabg5561, 2022.

[46] Carlo Pizzinelli, Augustus J Panton, Ms Marina Mendes Tavares, Mauro Cazzaniga, and Longji Li. *Labor market exposure to AI: Cross-country differences and distributional implications*. International Monetary Fund, 2023.

[47] David Saxton, Edward Grefenstette, Felix Hill, and Pushmeet Kohli. Analysing mathematical reasoning abilities of neural models. *arXiv preprint arXiv:1904.01557*, 2019.

[48] Manuel Trajtenberg. Ai as the next gpt: a political-economy perspective. Technical report, National Bureau of Economic Research, 2018.

[49] Michelle Vaccaro, Abdullah Almaatouq, and Thomas Malone. When combinations of humans and ai are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, pages 1–11, 2024.

[50] Michael Webb. The impact of artificial intelligence on the labor market. *Available at SSRN 3482150*, 2019.

[51] Darrell M West. *The future of work: Robots, AI, and automation*. Brookings Institution Press, 2018.