



Brief

Closing the Task Data Gap in Federal Statistics

Building New Measures for the AI Economy

Federal statistics name the tasks that make up each occupation, but no official data show what share of a job's time or output each task accounts for.

Without task-level weights and requirements, AI exposure studies compare model capabilities against task names alone, producing widely varying estimates of AI's effects.

A federally supported task-classification taxonomy, linked to Standard Occupational Classification (SOC) and NAICS codes and the BEA-BLS production accounts, would let agencies trace productivity changes to specific tasks.

William Martin
October 2026

The US statistical system faces a new challenge as firms and workers adopt large language models (LLMs): It does not yet collect the data required to understand how AI affects productivity. AI fundamentally changes work at the task level, so measuring the workplace effects of AI requires data on how tasks change because of AI and associated robotics. This data gap in official statistics ties directly into America’s AI Action Plan, which calls on federal statistical agencies to develop data collection methods to support the study of AI’s economic impact.¹ Firms are measuring AI’s effects; the government is only getting started.

Proposal

To close this gap in government statistics, we recommend the establishment of a federally supported task-classification taxonomy that connects to existing classification systems. Statistical agencies could build such a taxonomy through new surveys, through amendments to existing surveys, or through blending privately and publicly sourced data. A system that tracks each task’s requirements, economic weight, and other relevant data would provide the task-level information required to measure economic change more accurately.

Why It Matters

Current task-focused systems list the tasks that make up an occupation, but they cannot measure what share of a job’s output each task accounts for. Data on specific tasks also tend to omit physical, cognitive, and social requirements, the legal and ethical restrictions, and any other factors that might dictate a firm’s willingness to use AI.

The current statistical system is best described as task labeling rather than true measurement. Tasks are named and sometimes briefly described, but they are not linked to any datasets that might shed light on these factors. In other words, they provide information that is not easily converted into data used in cardinal measurements. As a result, many AI exposure studies operate by asking humans or LLMs to compare model capabilities against task names alone. This leads to a high level of variability in their exposure findings, as well as a lack of confidence in their measures. Without measuring what portion of a job’s output comes from each of its component tasks, we cannot measure how differences at the task level alter the contents of jobs, the resulting output of firms, and the productivity of those key drivers of national economic activity.

¹ The White House, “[Winning the Race: America’s AI Action Plan](#),” July 2025.

Nearly every question about AI is a question of magnitude; researchers want to know how much an occupation is exposed, or how much productivity changes when firms adopt new technology. Labels without meaningful measurement can only show that exposure or productivity gains exist, not their magnitude. This problem has received a broad range of interest from leading economic programs at Anthropic, OpenAI, Microsoft, Indeed, and the Stanford Digital Economy Lab.²

This uncertainty also casts doubt on other economic statistics affected by AI-driven productivity changes. Policymakers and businesses alike are largely unequipped to understand, through the lens of official statistics, how AI is affecting the labor market, which risks slower and poorer decisions.

Current systems, like the Occupational Information Network (O*NET) or the Occupational Requirements Survey (ORS), hold the building blocks for a task classification system. O*NET tracks the details of an occupation's component tasks, and ORS tracks the physical and cognitive requirements of a job. The problem is that most supplementary data (such as requirements) connect at the job level rather than the task level. For example, O*NET publishes task statements for roughly 900 occupations, but many relevant datasets link only to the occupations, not the tasks.

Worked Example: Construction

To further illustrate the benefits of task classification, take a task that O*NET lists for construction laborers: “Operate jackhammers or drills to break up concrete or pavement.”³ O*NET attaches three ratings representing the importance of the task, the share of workers who do the task, and the frequency (not duration) of the task. None of these ratings can be summed or apportioned to show how the job's time or output divides across its tasks.

Another source, the Occupational Outlook Handbook (OOH), which describes the work of occupations, lists six duties in its entry for construction laborers, much condensed from

² See, for example, Ruth Appel et al., *Anthropic Economic Index Report: Economic Primitives* (Anthropic, January 15, 2026); Tejal Patwardhan et al., “GDPval: Evaluating AI Model Performance on Real-World Economically Valuable Tasks,” OpenAI, September 25, 2025; Kiran Tomlinson et al., “Working with AI: Measuring the Applicability of Generative AI to Occupations,” Microsoft Research, arXiv:2507.07935, last revised December 22, 2025; Annina Hering and Arcenis Rojas, *AI at Work Report 2025: How GenAI Is Rewiring the DNA of Jobs* (Indeed Hiring Lab, September 23, 2025); Erik Brynjolfsson, Bharat Chandar, and Ruyu Chen, “Canaries in the Coal Mine? Six Facts About the Recent Employment Effects of Artificial Intelligence,” Stanford Digital Economy Lab, revised August 2026.

³ O*NET OnLine, “[Construction Laborers](#) 47-2061.00,” National Center for O*NET Development, updated 2026.

O*NET’s 27 tasks.⁴ The quantitative data on the page, such as median pay of \$47,120 (May 2025), employment of 1,489,300 (2025), or growth of 7 percent (2025–35), all describe the occupation itself. The duties appear only as brief statements of what a typical job might entail. OOH communicates job outlook statistics well, but its task duties carry no data.

Wage and requirement data also connect at the occupation level. For example, the Occupational Requirements Survey finds that construction laborers spend 89.2 percent of the workday standing.⁵ ORS does not state what share of that standing is jackhammering rather than digging, or which tasks require standing at all. Similarly, the American Time Use Survey records minutes worked by occupation, but not by task.⁶

At the industry level, the Bureau of Labor Statistics (BLS) publishes labor productivity for only four construction industries, which together account for about 12.5 percent of the sector’s employment.⁷ The rest of the sector has hours and employment data, but no productivity measures.

True measurement becomes scarcer at the task level. No statistic currently reports what share of a laborer’s time or output is attributable to the task of operating a jackhammer. If a new machine automated that task, its effect on the occupation and its output would be unknown.

Productivity

Many economists believe that AI will produce historic gains in productivity and economic growth, and some believe these gains are occurring now.

The current situation recalls the early 1990s, when the gains from desktop computing had not yet reached the labor statistics (the so-called Solow paradox). While the recession ended in March 1991,⁸ unemployment climbed to 7.8 percent in June 1992,⁹ and roughly eight in 10

⁴ US Bureau of Labor Statistics, “[Construction Laborers and Helpers](#),” Occupational Outlook Handbook, last modified August 27, 2026. Figures cited are for the detailed occupation of construction laborers (SOC 47-2061).

⁵ US Bureau of Labor Statistics, “[Construction Laborers](#),” Occupational Requirements Survey occupational profile, 2025 data.

⁶ US Bureau of Labor Statistics, [American Time Use Survey](#), accessed October 9, 2026.

⁷ US Bureau of Labor Statistics, “[Construction Labor Productivity](#),” updated September 24, 2025.

⁸ National Bureau of Economic Research, “[US Business Cycle Expansions and Contractions](#),” last updated March 14, 2023.

⁹ US Bureau of Labor Statistics, [Current Population Survey](#), unemployment rate, seasonally adjusted (series LNS14000000), accessed October 9th, 2026.

Americans still told Gallup the country was in recession.¹⁰ The Department of Commerce’s 1993 revision later found the recession “not as deep and the 1992 recovery not as weak as have been previously reported.”¹¹

Like then, the US economy may now be growing faster than measured because we lack the requisite data and metrics for tracking the change. Official productivity statistics report output per hour for industries and sectors, but they cannot show which tasks within an occupation are subject to AI or robotics-driven change, or by how much. A task-based system that records requirements and economic weights would allow statistical agencies to trace changes in task performance upward through occupations and industries to better measure productivity. It would also help fulfill the AI Action Plan’s directive that BLS, the Census Bureau, and the Bureau of Economic Analysis (BEA) study AI’s impact on the labor market using data they already collect.¹²

Integrating task data into productivity accounts through a supplementary table could add a task dimension to the BEA-BLS Integrated Industry-Level Production Account.¹³ Task-level efficiency gains are reflected in this account at the industry level but are not attributed to the specific task-level contributions. If fewer hours produce the same output, or if output rises while hours remain steady, the gain simply appears as higher labor productivity or total factor productivity (TFP). Either way, the task-level change that generated this efficiency gain remains invisible. A supplemental task table would add a layer below the industry level. Task time-shares can be aggregated to occupation hours, and occupation hours would need to be bridged to industry hours.

The real challenge lies in bridging occupational hours to industry hours. This bridge would need to be developed because the Occupational Employment and Wage Statistics (OEWS) program reports occupational employment by industry but not hours, and the BEA-BLS Integrated Industry-level Production Accounts do not have the necessary occupational details. Once the bridge is in place, analysts could trace changes in an industry’s labor productivity or TFP to specific tasks. If a new machine cut the time construction laborers spend breaking up concrete, we could show how much of the sector’s productivity change came from that one task.

¹⁰ Gallup, “[Most Americans Say U.S. Is Now in Economic Recession](#),” March 19, 2008, reporting that between 79 percent and 84 percent of Americans said the economy was in a recession in three Gallup polls conducted in 1991 and 1992.

¹¹ US Department of Commerce, Bureau of Economic Analysis, “[Annual Revision of the U.S. National Income and Product Accounts](#),” *Survey of Current Business* 73, no. 8 (August 1993).

¹² The White House, “Winning the Race: America’s AI Action Plan.”

¹³ US Bureau of Economic Analysis, “[Integrated Industry-Level Production Account \(KLEMS\)](#),” accessed October 9, 2026.

What Would Change

Task-level productivity measurement does not currently exist in a form sufficient to support sound policy decisions during an AI transition. No single federal dataset supports a full analysis of tasks. O*NET describes tasks without their requirements or economic importance, and ORS measures requirements without reference to tasks. New or amended surveys, linked to SOC and NAICS codes and the BEA-BLS production accounts, could remedy this. Such a system would fill an evident gap in official statistics and give both policymakers and businesses faster, more accurate productivity statistics.

The Fiscal Lab is now working on this project in coordination with federal agencies. We are developing an in-depth research paper on the current state of task-level data across O*NET, ORS, OEWS, and other surveys, and on possible ways to address the gap with private and public data. We are also in conversations with the Office of Management and Budget to coordinate the development of this system. Historically, our understanding of the economic implications of previous technologies suggests that the gains from AI may only be found well after the fact. Yet we may still have an opportunity to build the right measurement system in time to understand the effects of this technology at the task level and better inform the decisions of public and private stakeholders. If you are working on similar research or would like to get involved, please reach out to the Fiscal Lab.

William Martin is an intern at the Fiscal Lab on Capitol Hill.