2026 No. 051

Indexing the Impact of AI within the O*NET System: A Review of Methods and Development of Recommendations

Prepared for: National Center for O*NET Development
313 Chapanoke Road, Suite 130
Raleigh, NC 27603

Prepared under: Subcontract Number
(through RTI International):
1-312-0207142--41224L

Authors: Dan J. Putka, HumRRO
Nathaniel M. Voss, HumRRO
Phil Lewis, National Center for O*NET Development

Date: June 18, 2026

Indexing the Impact of AI within the O*NET System: A Review of Methods and Development of Recommendations

Table of Contents

Table of Contents (Continued)

List of Tables

List of Figures

Indexing the Impact of AI within the O*NET System: A Review of Methods and Development of Recommendations

Introduction

This report provides a review of methods for indexing the “impact” of Artificial Intelligence (AI) on work at scale, along with research-based recommendations for developing a suite of AI impact indices for the Occupational Information Network (O*NET). Specifically, our first focus is on reviewing the methods past studies have used to evaluate how AI can influence work in occupations. Over the past 5 years, there has been a rapid proliferation of research and popular press focused on AI’s impact on work, including the potential for AI to automate job performance and augment humans’ performance of various job tasks. For the most part this research has arisen out of domains of labor economics and computer science, with less attention arising from the domain of Industrial-Organizational (I-O) psychology wherein the scientific study of job performance, as well as jobs’ performance requirements, have been grounded for the past 100 years (e.g., Brannick et al, 2007; Campbell & Wiernik, 2015; Gael, 1988; McCormick 1979; Wilson, 2007; Wilson et al., 2012).

Our focus here is on methods that *scale* well – specifically, those that can be used efficiently to draw inferences about AI’s impact on work across an entire population of occupations rather than a limited set. Perhaps not surprisingly, an extensive amount of the research we review leveraged the occupational taxonomy and data underlying the U.S. Department of Labor’s (DOL) O*NET System (e.g., Appel et al., 2026; Brynjolfsson et al., 2018; Eloundou et al., 2024; Felten et al., 2021; Hampole et al., 2025; Handa et al., 2025; Kochhar, 2023; Shao et al., 2025; Tomlinson et al., 2025; Webb, 2019). Indeed, of the 19 core studies we reviewed for this report, 16 of these (84%) used O*NET data in some manner.[1] O*NET is a comprehensive system developed and maintained by the DOL that provides information for over 900 occupations, encompassing over 57,000 jobs within the U.S. economy. Occupational information is maintained in a comprehensive database (National Center for O*NET Development, 2026a). To keep the database current, the National Center for O*NET Development (hereinafter referred to as “the Program”) conducts a continuous data collection process to identify and maintain current information on the characteristics of workers and occupations. As such, it provides a meaningful and well-structured starting point for studies of AI’s impact on work.

Purpose

The purpose of this review is to develop research-based recommendations for a set of indices that the O*NET Program could develop to index the impact of AI on skills and occupations included in O*NET. The goal is not to propose the Program “reinvent the wheel”, but rather draw on work performed to date, while at the same time being sensitive to the unique needs and nuances of maintaining such data for O*NET (e.g., minimizing dependencies on external, non-government data sources that have no guarantee of being maintained or refreshed over time). The recommendations we offer to the Program for developing a suite of AI impact indices for use within O*NET are informed not only by observations from past AI impact research but also grounded in the historical and modern job analysis and job performance literatures.

[1] In the *References* section of this report, we identify specific studies that used O*NET data.

Overview of Report Structure

In the remaining sections of the report, we provide a methods-focused review of the AI impact literature and our recommendations for O*NET. We begin with a summary of our review strategy and the development of an "AI Impact Methods Framework" to help organize our review. We developed this framework early in the review process to facilitate discussion of methods. Effectively, the framework allows one to quickly organize and categorize methods studies have used to index the impact of AI on work. We organize our review in terms of key elements of this framework, namely:

- **AI Conceptualization**: To understand AI's impact on work, one must first have an operational definition of "AI". As such, we first review and summarize the different ways studies have operationalized AI to examine its impact on work.
- **AI Impact Definition/Scaling**: Beyond the differences in how AI is operationalized, studies also differ in terms of how they define and scale AI "impact". Here, we use the term "impact" loosely, realizing a major differentiator of studies is how they define or scale impact (e.g., in terms of AI *automating* work, in terms of AI *augmenting* humans' ability to perform work, or in terms of an occupation's *exposure* to AI more generally).
- **Drawing Inferences About AI's Impact on Specific Elements of Work**: Once the operationalization and scaling of AI have been specified, researchers typically focus assessing AI's impact on a specific type of work element that falls below the occupation-level (e.g., linking capabilities of AI to the knowledges, skills, and abilities need to perform a job effectively, assessing AI's impact on tasks performed on a job). As such, we review and summarize different ways studies have indexed AI's impact on elements of work.
- **Drawing Inferences about AI's Impact on Occupations:** Finally, many studies of AI impact have aggregated AI impact scores computed for the various work elements to the occupation-level to derive an occupation-level estimate of AI impact. Accordingly, we also review and summarize the methods used to assess AI's impact at the occupational level.

After reviewing methods for indexing AI impact used in past studies, we transition to detailing our recommendations for developing a suite of AI impact indices for O*NET. We begin by discussing key considerations and observations about past AI impact work that influenced our recommendations. Next, we describe each proposed index, its rationale, and how it could be used in the context of O*NET and benefit O*NET users. We conclude our discussion with initial thoughts on how the O*NET Program may evaluate such indices, and potential future possibilities it may want to consider should it move forward with developing the recommended indices.

AI Impact Methods Literature Review

Review Strategy

For our literature review, we aimed to identify studies that examined the impact of AI on occupations or pertinent work elements, such as tasks and Knowledge, Skills, Abilities, and/or Other Characteristics (KSAOs). To help contextualize these findings and inform our recommendations, we also reviewed various AI frameworks, definitions, and taxonomies. For our search, we used common databases such as Google Scholar and EBSCOhost. We also used basic internet searches, since certain sources may not always appear in research databases (e.g., reports from research organizations, think tanks, non-profits). Given the breadth of this topic, we used a wide variety of terms to identify relevant research studies. This included, but was not limited to, terms such as “AI occupation impact”, “AI economic impact”, “AI occupation exposure”, “AI task automation”, “AI task augmentation”, “AI usage at work”, “Skills automated by AI”, and “Skills augmented by AI”. Note that we were primarily interested in the effects of AI on tasks, KSAOs, occupations, etc., and less so in studies exploring economic outcomes (e.g., employment, wages). For each article we identified, we supplemented our search through snowballing methods by reviewing both the article’s reference list (backward snowballing) and articles that cited it (forward snowballing). We also employed Large Language Models (LLMs) to search the web for relevant references while regularly inspecting results to ensure accuracy. Given the rapid pace at which AI impact studies are currently being published, we conducted our search on an ongoing basis throughout the review to ensure that recently published studies were included.

Formulation of an AI Impact Methods Framework

In the early stages of our literature review, we began to see ways to classify the methods AI impact studies used to draw inferences regarding AI’s impact on work. Specifically, early in our process, we identified key differentiating features of studies that, when combined, could help classify the approaches they adopted to assess AI impact. These key differentiating features included (a) how AI was operationalized, (b) the initial target of the AI impact measure (e.g., often this was individual tasks, skills, job vacancies rather than occupations directly), and (c) the method used to measure AI impact at the occupation-level (e.g., often this involved aggregating a task- or skill-level AI impact measure to the occupation-level).

In light of these early observations, we developed a conceptual framework to aid in organizing our AI impact methods review, which we used to categorize the majority of studies we reviewed by the combination of AI operationalization, initial target of AI impact, and occupation-level aggregation strategy they adopted. Figure 1 illustrates this framework, and in the sections that follow, we describe where various past AI impact studies fall with respect to the key features that define it.

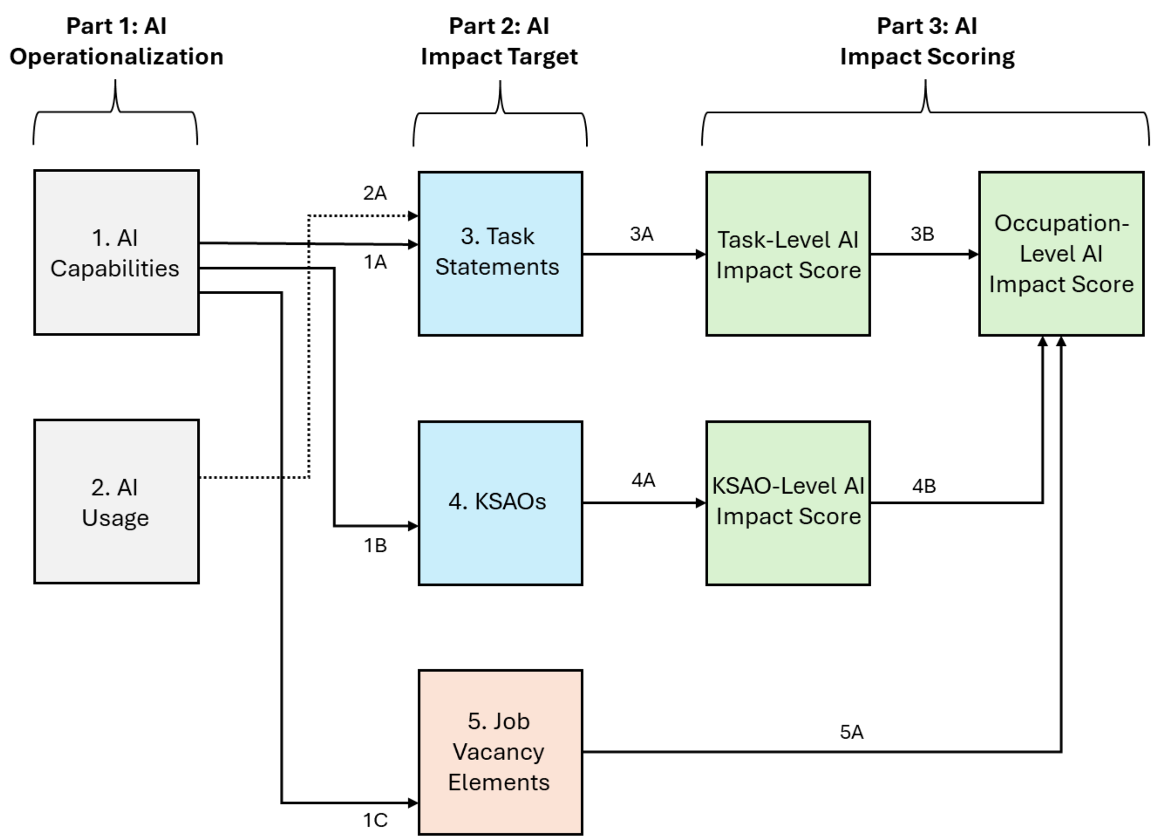
In short, most studies begin with a discussion about their operationalization of AI. We have broken these operationalizations into two broad categories. The first focused on what we organize under “AI capabilities” (Figure 1, Part 1.1), which may manifest as:

- A general definition of a broad category of AI of focus in the study (e.g., generative AI, machine learning).
- A listing of the domains in which AI functions of the general types of capabilities AI may exhibit (e.g., Language, Social Interaction, Vision, Robotic Intelligence).
- A listing of specific AI-related technologies, applications, methods/techniques of interest in the study.

The second broad category focuses on "AI usage" (Figure 1, Part 1.2), operationalizing AI through usage patterns associated with specific Generative AI (GenAI)-powered applications (e.g., Microsoft Copilot, Anthropic's Claude platform).

With an operationalization of AI specified, studies often have an initial target for their AI impact question of interest. Initial targets are often specific elements of work, such as job tasks or job-relevant KSAOs, though some studies focus more generally on various elements of job vacancies (Figure 1, Paths 1A, 1B, 1C, 2A). The reason why we refer to these as "initial targets" is that often times studies typically use the initial task-level or KSAO-level AI impact scores they form as an intermediate step via which they draw inferences about AI's impact on occupations (e.g., by aggregating task or KSAO-level AI impact scores across tasks/KSAOs within an occupation) (Figure 1, Paths 3A-3B, 4A-4B).

Figure 1. AI Impact Literature Review Framework



Though this framework can be used to classify most past AI impact studies by the combination of "paths" they take to get from an operationalization of AI to inferences about AI's impact on occupations, there are some exceptions. For example, some studies skip making intermediate inferences about AI impact on work elements such as job task or skills, and inquire directly about AI impact on an occupation in general (e.g., Frey & Osborne, 2017) or stop short of assessing AI impact on occupations, and only focus on assessing impact at the work element level (e.g., job tasks, see Cai et al., 2026). Moreover, the lack of certain paths in Figure 1 (e.g., AI usage-to-KSAOs, AI usage-to-vacancy descriptions) is not meant to imply these paths do not exist, merely that they are less common within the current literature. Thus, missing paths may indicate strategies that have yet to be explored or are underexplored in the literature to date. Lastly, while it is possible to adopt more than one path, we rarely observe this in a given study (with one notable exception being Garcia-Lluis Valencia's (2026) multi-dimensional "ITEA Framework" for assessing automation exposure to agentic AI, which employs several work elements and numerous AI impact indicators). In sum, for most AI impact studies to date, Figure 1 can be used to classify studies by the configurations of elements and paths they adopted to move from an initial conceptualization of AI to an inference about AI's impact on skills and occupations. Because of this, the remainder of our review is organized around this framework.

Operationalizing AI for Purposes of Studying Its Impact on Work

As noted above, to understand AI's impact on occupations, one must start with an operational definition of AI.[2] A summary of the AI operationalizations for a set of exemplar studies we reviewed is provided in Table 1. Note that these 19 studies also served as a set of core studies we reviewed for cataloging methods to assess AI impact, and we map back to Figure 1 in the sections that follow. This is not meant to imply that these are the only studies that have examined this topic, though we believe they are representative with regard to methods for assessing AI impact and can provide a foundation for understanding how AI impact has been explored to date. Moreover, many of these studies have been widely cited, and thus reviewing them is critical for understanding prominent ways AI impact has been conceptualized (particularly since these studies often serve as a baseline for comparing results from newer studies). Additionally, several of the studies we reviewed were published very recently, and their inclusion helps ensure that a wide range of cutting-edge methodologies (e.g., LLM-based estimation methods) were considered in our review. In sum, the studies we examined in depth provide a representative foundation for which to understand methods used to assess AI impact.

The majority of studies examining AI's impact tend to operationalize AI either through its capabilities (Figure 1, Part 1.1) or through the ways people actually use generative AI-powered applications (Figure 1, Part 1.2). With regard to capabilities, both Brynjolfsson et al. (2018) and Webb (2019) adopt algorithm-based conceptualizations of AI focused on tasks that Machine Learning (ML) methods can perform. Other conceptualizations of AI are a bit broader in their framing of AI, such as Organisation for Economic Co-operation and Development's (OECD's) (2025) AI Capability Indicators, which benchmark AI capabilities across nine domains (e.g., language, social interaction, problem solving) and Hampole et al.'s (2025) 20 AI application categories (e.g., task and workflow automation, text and knowledge retrieval). In other cases, AI is operationalized more narrowly, as in Ahmadi et al.'s (2024) focus on ChatGPT skills specified in job advertisements. Despite employing unique operationalizations of AI, the common theme underlying most research on AI capabilities is an evaluation of how AI (however defined and operationalized) may likely impact occupations, tasks, KSAOs, or other work elements.

In contrast, other studies explore the impact of AI by observing which real-world activities individuals are actually using generative AI-powered applications to perform. These studies typically operationalize AI by analyzing millions of real-world interactions with Microsoft Bing's Copilot or Anthropic's Claude platforms to identify how generative AI is being used for specific job tasks. That is, rather than relying on theoretical judgments (i.e., asking a human/LLM what job tasks AI can potentially augment and/or automate), such studies map anonymized conversations or application usage to O*NET occupational information, such as ~20,000 O*NET

[2] As part of our review, we also examined high-level AI taxonomies, definitions, and frameworks that have been published in recent years (see Appendix A). We do not focus on these in the body of the report because they are somewhat tangential to our focus, which is on how specific studies have operationalized AI for assessing its impact on work.

task statements or 332 O*NET Intermediate Work Activities (IWAs), to observe which tasks/IWAs these generative AI-powered applications are being used to inform or perform (e.g., Appel et al., 2026; Handa et al., 2025; Tomlinson et al., 2025).

Table 1. Summary of How AI has been Operationalized Across Select AI Impact Studies

Reference	AI Focus	AI Operationalization
Agarwal et al. (2025)	Usage	AI is operationalized as 35 unique "parameters" (e.g., information processing, complexity of choice) grouped into seven core categories (e.g., social intelligence, decision-making).
Ahmadi et al. (2024)	Capabilities	AI is operationalized as market demand for specific ChatGPT-related skill sets in job announcements.
Alekseeva et al. (2021)	Capabilities	AI is operationalized as a list of 71 AI skill requirements (e.g., AI chatbot, machine learning, computer vision) included in vacancy announcements.
Appel et al. (2026)	Usage	AI is operationalized as different types of tasks users perform with a specific generative AI application, Anthropic's Claude platform. This is integrated with five "economic primitives" (e.g., task complexity and the degree of autonomy delegated to Claude), which provide nuanced insights into how AI is used for specific tasks.
Brynjolfsson et al. (2018)	Capabilities	AI is operationalized as Machine Learning (ML), a subfield of AI, defined as computer programs that automatically improve their performance at tasks through experience.
Chopra et al. (2025)	Capabilities	AI is operationalized as a set of ~13,000 tools mapped to a standardized skill taxonomy.
Eloundou et al. (2024)	Capabilities	AI is operationalized as a GPT or GPT-powered system.
Felten et al. (2021)	Capabilities	AI is operationalized as 10 AI applications (e.g., language modeling, speech recognition) as measured by the Electronic Frontier Foundation (EFF).
Fossen & Sorgner (2019)	Capabilities	AI is operationalized via three metrics with a different AI impact focus that have been developed elsewhere (i.e., Frey & Osborne's [2017] index, Felten et al.'s [2018] index, and Brynjolfsson et al.'s [2018] index).
Hampole et al. (2025)	Capabilities	AI is operationalized as resume-derived firm-year AI use cases, identified by screening "AI-integrator" resumes with ~30 AI-related keywords and then using an LLM to extract the specific AI applications those workers reported developing or deploying.
Handa et al. (2025)	Usage	AI is operationalized as different types of tasks users perform (e.g., designing information technology systems, customer service operations) with a specific generative AI application, Anthropic's Claude platform.
Kochhar (2023)	Capabilities	AI is operationalized using a general definition that describes it as applications (e.g., ML, computer vision, and Natural Language Processing [NLP]) and tools (e.g., ChatGPT) that allow machines to either substitute for or complement human activities.

Table 1. (Continued)

Reference	AI Focus	AI Operationalization
Mäkelä & Stephany (2025)	Capabilities	AI is operationalized as a list of 115 unique skills, classified into eight categories: "AI Creators" (e.g., reinforcement learning) and "AI Users" (e.g., ChatGPT, sentiment analysis).
Organisation for Economic Co-operation and Development (OECD) (2025)	Capabilities	AI is operationalized as nine different AI capability areas (e.g., language, social interaction, vision, robotic intelligence).
PricewaterhouseCoopers (PwC) (2025)	Capabilities	AI is operationalized the same way as Felten et al. (2021) but also supplemented with the International Monetary Fund (IMF) Complementarity Index (Pizzinelli et al., 2023). In a separate analysis, AI was operationalized as a list of 376 AI skills.
Shao et al. (2025)	Capabilities	AI is operationalized as "AI agents", which are systems capable of autonomously designing workflows and using software tools to perform multi-step tasks on behalf of a user.
Sigelman et al. (2026)	Capabilities	AI is operationalized as AI-exposed skills present within job postings.
Tomlinson et al. (2025)	Usage	AI is operationalized in terms of different ways users interact with a specific generative AI application, Microsoft Bing Copilot, with a key distinction between the "User Goal" (assistance sought) and the "AI Action" (task performed).
Webb (2019)	Capabilities	AI is operationalized through ML algorithms (specifically supervised and reinforcement ML).

Drawing Inferences About AI's Impact on Specific Elements of Work

Upon reviewing and summarizing how AI was operationalized within studies exploring AI impact, we next sought to understand how past work has made inferences about AI's impact on specific work elements (e.g., job tasks or skills, elements of job vacancies; Figure 1, Paths 1A, 1B, 1C, 2A). The methods used to create AI impact scores at the task, KSAO, and vacancy element level are summarized in Table 2 and also described in greater detail below.

A key differentiating factor of the studies in Table 2 is how they have defined "impact" (see Impact Type in Table 2). One common way impact has been defined is in terms of general AI *exposure*, without necessarily making an explicit distinction between AI automating and augmenting tasks, skills, or other work elements. For example, Kochhar's (2023) study of AI exposure explores whether the "activities workers perform on their jobs may be replaced or aided by artificial intelligence." (page 2). Moreover, Felten et al. (2021), with regard to their measure of occupational exposure to AI, are more explicit when they state, "we are agnostic about whether or when AI will augment or replace human labor." (page 2202).

In contrast to a focus on general AI *exposure*, other studies focus on the distinction between AI *automation* and AI *augmentation*, such as:

- PricewaterhouseCooper's (PwC's) (2025) study of AI impact that employed Felten et al.'s (2021) index along with the International Monetary Fund's (IMF) classification of augmentable and automatable jobs (Pizzinelli et al., 2023).
- Sigelman et al.'s (2026) complementary view of AI automation/augmentation (see also Chen et al.'s (2025) distinction between AI displacement and complementarity).
- Shao et al.'s (2025) Human Agency Scale, which explicitly quantifies the degree of human involvement that is needed to complete tasks that can also be completed with AI.
- Handa et al.'s (2025) and Appel et al.'s (2026) examination of real-world AI usage patterns to distinguish between tasks being automated by AI and tasks where AI is used for augmentation purposes.

Once the type of AI impact is established (e.g., exposure, automation, augmentation), studies then compute AI impact estimates for specific work elements (see Table 2). In the sections that follow, we organize our review of studies by whether the study targeted job tasks, job-relevant KSAs, or job vacancies.

Table 2. Summary of Methods for Drawing Inferences About AI's Impact on Elements of Work

Reference	Target	Impact Type	Method Summary
Agarwal et al. (2025)	Task	Augmentation	Used Gemini 2.5 Pro to score ~3,500 O*NET tasks on a 1-10 scale across 35 parameters (e.g., idea generation, routineness). These were averaged into 7 primary characteristic scores (e.g., cognitive, social, creativity). Next, used principal components analysis (PCA) and k-means clustering to group tasks into empirically derived "task archetypes" based on underlying characteristics (e.g., cognitive demand, creativity, routineness). AI impact was assessed by comparing the frequency of AI usage across these task archetypes.
Appel et al. (2026)	Task	Automation and Augmentation	AI impact was computed as "effective AI coverage" for O*NET tasks (exact number of tasks explored unspecified), which represents the weighted sum of task-specific success rates, where each task's weight is defined as the share of the worker's time (adjusted by how frequently the task occurs), which draws on Claude success rate data, hours estimates from prior productivity work, and O*NET frequency data.
Brynjolfsson et al. (2018)	Task	Automation	2,069 O*NET Detailed Work Activities (DWAs) were rated by humans on a multi-item rubric (23 criteria, 1-5 scale) capturing whether ML can learn the input–output mapping for that task (i.e., the "Suitability for Machine Learning" metric).

Table 2. (Continued)

Reference	Target	Impact Type	Method Summary
Eloundou et al.(2024)	Task	Exposure	2,087 O*NET DWAs and 19,265 O*NET tasks were evaluated using a rubric that links task requirements to GPT capabilities, with tasks classified as “exposed” if GPTs could substantially reduce completion time by at least 50%. Exposure ratings were generated using both human annotators (for DWAs and a subset of tasks) and GPT-4 (for all tasks) to distinguish between tasks GPT can perform independently (direct exposure) and those requiring additional tools or software.
Hampole et al. (2025)	Task	Exposure	About 20,000 O*NET tasks and 20 AI applications are embedded and compared via cosine similarity. An application is considered relevant to a task if similarity exceeds the 95th percentile, and task-level exposure is then defined as the share of a firm’s AI applications that meet this criterion.
Handa et al. (2025)	Task	Automation and Augmentation	AI impact at the task level was operationalized as the observed frequency with which real-world AI interactions (via Claude conversations) are classified into specific O*NET tasks using a hierarchical tree mapping procedure.
Kochhar (2023)	Task	Exposure	Human raters categorized 41 O*NET General Work Activities (GWAs) into Low, Medium, or High AI exposure based on whether AI could perform or assist them. Consensus was reached by reviewing the detailed sub-tasks for ambiguous activities.
OECD (2025)*	Task	Exposure	See Table note.
Shao et al. (2025)	Task	Automation and Augmentation	Developed the Human Agency Scale (HAS), where workers rated their “automation desire” for 844 O*NET tasks (1-5 Likert scale), while AI experts independently rated the current technological feasibility for the same tasks using the HAS levels.
Tomlinson et al. (2025)	Task	Automation and Augmentation	Assessed AI impact by mapping real-world AI interactions via Copilot conversations to 332 O*NET IWAs, doing so separately for user goals (what the user is trying to accomplish, reflecting AI augmentation) and AI actions (what the AI is actually performing). For each IWA, measured whether AI is used non-trivially, how successfully AI performs/assists the task, and how broadly AI contributes to the task.

Table 2. (Continued)

Reference	Target	Impact Type	Method Summary
Webb et al. (2019)	Task	Automation	Extracted verb-noun pairs (e.g., "diagnose disease", "recognize image") from AI patent titles using natural language processing algorithms and mapped these to verb-noun pairs extracted from O*NET tasks. Each occupational task is then scored based on how frequently similar pairs appear in AI patents.
Chopra et al. (2025)	KSAO	Automation	Mapped over 13,000 AI tools to a taxonomy of over 32,000 skills. Large population models were used to simulate whether AI can technically perform each skill based on tool descriptions and metadata.
Felten et al. (2021)	KSAO	Exposure	Assessed AI exposure by using crowd-sourced ratings to link 10 AI applications to 52 O*NET abilities, then summing application-ability relatedness scores to estimate how strongly each ability was exposed to contemporary AI capabilities.
Fossen & Sorgner (2019)**	KSAO and Task	Multiple**	Assessed the impact of new digital technologies using three occupation-level measures that have been previously developed elsewhere: (1) computerization risk (Frey & Osborne, 2017), (2) advances in AI (Felten et al., 2018), and (3) suitability for machine learning (Brynjolfsson et al., 2018).
PwC (2025)	KSAO	Automation and Augmentation	Combined Felten et al.'s (2021) Artificial Intelligence Occupational Exposure (AIOE) with the IMF Complementarity Index (Pizzinelli et al., 2023) to explicitly classify roles as either augmented or automated. Also tracked demand for AI skills over time based on a list of 376 AI skills.
Ahmadi et al. (2024)	Vacancy	Exposure	The study collected job postings mentioning "ChatGPT" and then used Latent Dirichlet Allocation (LDA) topic modeling on the ChatGPT-relevant text within those postings to assign each vacancy a probability score across five skill sets (e.g., prompt engineering, creative content generation).
Alekseeva et al. (2021)	Vacancy	Exposure	Identified "AI vacancies" by tagging job postings that required at least one AI-related skill (e.g., deep learning, speech recognition) from a list of skills based on the Burning Glass Technologies (BGT) skills taxonomy.

Table 2. (Continued)

Reference	Target	Impact Type	Method Summary
Mäkelä & Stephany (2025)	Vacancy	Automation and Augmentation	Assessed at the individual job posting level by using vacancy data to identify AI-related postings (via the presence of at least one AI skill) and then analyzing how the presence of complementary vs. substitutable skill requirements, and their associated wages, differed within and across these job postings
Sigelman et al. (2026)	Vacancy	Automation and Augmentation	Used job postings to identify skills within occupations, then applied a multi-LLM approach to score each skill-occupation pair on two continuous dimensions: (1) exposure to AI replacement (substitution) and (2) exposure to AI enhancement (augmentation).

Note. *While linkages to tasks have not yet been made (to our knowledge), OECD proposed linking the capability indicators to tasks by mapping each task's human ability requirements (e.g., from O*NET) onto the AI Capability Indicator scales, then rating the level of capability required for that task and comparing it to current AI capability levels to assess coverage or gaps. **Authors used multiple indices with different AI targets and impact types (i.e., Frey & Osborne's [2017] index, Felten et al.'s [2018] index, and Brynjolfsson et al.'s [2018] index). Target refers to the work element that the study investigated regarding AI impact. Impact Type refers to whether the researcher's AI impact metric(s) focus primarily on AI impact in a general sense (e.g., exposure) or if they made a more explicit distinction between AI automation and AI augmentation (or explored both of these).

Drawing Inferences About AI's Impact on Job Tasks

Most studies exploring AI impact tend to focus on job tasks (or task-like information such as O*NET's Detailed Work Activities [DWAs], Intermediate Work Activities [IWAs], and/or General Work Activities [GWAs]; Figure 1, Path 1A).

Several studies have operationalized AI in terms of capabilities and examined its impact on job tasks (Figure 1, Path 1A-3A). For example, Kochhar (2023) evaluated AI exposure by having human annotators rate the likelihood that human labor on O*NET GWAs could be substituted or complemented by AI. Additionally, Brynjolfsson et al. (2018) developed a "Suitability for Machine Learning" (SML) score for each O*NET DWA. Hampole et al. (2025) identified firm-specific AI applications from resumes of AI users within organizations (i.e., "AI integrators") and then used Natural Language Processing (NLP) techniques to calculate the semantic similarity between those applications and O*NET task statements. Furthermore, Eloundou et al. (2024) examined both O*NET tasks and DWAs and used a rubric in which human annotators and/or an LLM (i.e., GPT-4) evaluated whether AI would reduce the time it takes to complete a task/DWA by at least 50%.

Additionally, though less common, multiple studies have operationalized AI in terms of individuals' actual use of generative AI-powered applications and examined its impact on job tasks (Figure 1, Path 2A-3A). A common approach is to analyze LLM conversations and map each interaction to the O*NET task taxonomy. For example, Handa et al. (2025) classified millions of Claude conversations into O*NET task categories using an LLM-based method and operationalized task-level AI usage as the frequency with which tasks appear in real-world interactions. Likewise, Agarwal et al. (2025) linked observed AI usage to tasks using LLMs by assigning each O*NET task scores on dimensions such as complexity, creativity, routineness, and cognitive demand, and then examined how these features predicted AI engagement.

Similarly, Tomlinson et al. (2025) analyzed Microsoft Bing Copilot conversations by decomposing each interaction into a user goal (i.e., the task the human is trying to accomplish) and an AI action (i.e., the task the AI performs). This was then mapped to O*NET IWAs, whereby AI impact is computed using multiple approaches: (1) how often a task appears in AI interactions, (2) whether the AI successfully completes the task, and (3) the scope of AI contribution (i.e., how much of the task is performed by AI). Finally, Appel et al. (2026) further extended this approach to explore the impact of AI by introducing "economic primitives" (e.g., task complexity, success rates, and degree of autonomy) in the context of Claude conversations. These primitives explicitly quantify how AI performs on tasks (e.g., whether it succeeds, how complex the task is, and whether the interaction reflects automation or augmentation), thereby providing a more nuanced measure of how AI is used to complete specific job tasks.

Drawing Inferences About AI's Impact on Job-Relevant KSAOs

In addition to task-based approaches, a growing stream of research has examined AI's impact on job-relevant KSAs (Figure 1, Path 1B-4A) by directly linking AI capabilities to underlying human abilities and skills. One example is Chopra et al. (2025), which introduces the Iceberg Index, a skills-centered measure of technical automation exposure. In sum, this approach operates entirely at the skill level by mapping over 13,000 AI tools to a listing of 32,000+ skills and using large population models to simulate whether AI systems can perform each skill based on tool capabilities and metadata.[3] This produces a skill-level automatability signal, which is then weighted by the economic value (i.e., wage contribution) of each skill to quantify the proportion of skill value that is technically performable by AI. Similarly, Felten et al. (2021) developed an ability-level measure of AI exposure by explicitly linking AI applications to O*NET abilities. Using a crowd-sourced dataset, the authors estimated the relatedness between 10 core AI applications (e.g., language modeling, image recognition, speech recognition) and 52 O*NET abilities. These relationships are then aggregated to generate an ability-level exposure score that reflects the extent to which each human ability is associated with current AI capabilities.

Moreover, Fossen and Sorgner (2019) incorporated various indicators by drawing on three distinct measures of AI impact that were previously developed (i.e., Frey & Osborne's [2017] index, Felten et al.'s [2018] index, and Brynjolfsson et al.'s [2018] index) and which assess AI impact at multiple levels of analysis (i.e., tasks and KSAOs). Finally, other recent applied work, such as PwC's (2025) study, builds directly on these KSAO-level approaches by combining Felten et al.'s (2021) ability-based exposure index with the IMF's AI-human complementarity index, which enabled them to make a more explicit distinction between AI automation (substitution of skills) and augmentation (enhancement of skills). In this framework, KSAO-level exposure serves as the primary input, while the classification of impact type depends on whether AI capabilities replace or support those underlying human attributes.

[3] Note, Chopra et al (2025) cite O*NET and the Census' American Community Survey (ACS) as the source of the 32,000+ standardized skills referenced in their work, but neither of these sources offers a skills taxonomy at that level of granularity, so it is unclear what the source of skills was in this study.

Drawing Inferences About AI's Impact on Work via Job Vacancies

Regarding AI capabilities, a third methodological approach examines AI's impact through job vacancies (Figure 1, Path 1C). Rather than inferring AI's impact from tasks or KSAOs, these studies directly analyze employer demand for specific skills, treating vacancies as real-time indicators of how AI is reshaping the skills required. For example, Sigelman et al. (2026) employ a large-scale, LLM-based approach to classify and track skills mentioned in millions of job postings over a five-year period (i.e., pre- and post-ChatGPT). Specifically, they use LLMs to identify and categorize skills derived from job postings and then classify them based on their exposure to AI automation versus augmentation. This allowed the authors to observe how demand for different types of skills is changing over time. Similarly, Alekseeva et al. (2021) used a vacancy-based, data-driven approach to analyze online job postings to identify AI-related skill demand and adoption patterns. Their method used textual analysis of job descriptions to detect the presence and growth of AI-related skills, effectively using job descriptions as a proxy for firm-level adoption of AI technologies.

Other recent work has extended this research using other techniques. For instance, Ahmadi et al. (2024) apply text mining and topic modeling to job advertisements to extract and categorize distinct clusters of generative AI-related skills, such as prompt engineering, content generation, and product development. Finally, Mäkelä and Stephany (2025) examined approximately 12 million job vacancies to examine how AI reshapes skill demand through substitution and complementarity. The researchers operationalized "AI-related skills" endogenously from the vacancy data itself by identifying postings that explicitly require AI capabilities (e.g., roles mentioning AI tools, methods, or technical competencies) and treating these as "AI-intensive" roles. Rather than relying on predefined taxonomies, this approach uses the presence of AI requirements within job postings to define AI exposure.

Drawing Inferences About AI's Impact on Occupations

Once AI inferences are made for specific work elements (e.g., tasks, KSAOs), studies have often aggregated element-level scores up to the occupation-level to make inferences about AI's impact on occupations (Figure 1, Paths 3B, 4B, 5A). In some cases, AI impact scores are based solely on vacancy-level descriptions and then used to directly produce occupation AI impact scores (Figure 1, Path 5A). In most cases, however, AI impact scores are first computed for either tasks (Figure 1, Path 3A) or KSAOs (Figure 1, Path 4A) and then aggregated to the occupation-level using different methods (Figure 1, Paths 3B and 4B). Below, we describe in greater detail some representative approaches for generating AI impact scores, with a summary of methods for assessing AI impact at the occupational level provided in Table 3.

Table 3. Summary of Methods for Drawing Inferences AI's Impact on Occupations

Reference	Target	Impact Type	Method Summary
Agarwal et al. (2025)*	Task	Augmentation	See Table note.
Appel et al. (2026)	Task	Automation and Augmentation	Occupational AI impact indices were calculated as weighted sums of O*NET task-level success rates, with each task weighted by its share of a worker's time (adjusted for task frequency).
Brynjolfsson et al. (2018)	Task	Automation	Aggregated (importance-weighted) O*NET DWA-level SML scores to compute a mean occupation-level measure of ML suitability.

Table 3. (Continued)

Reference	Target	Impact Type	Method Summary
Eloundou et al. (2024)	Task	Exposure	Human annotators provided labels at the O*NET DWA level. Since tasks can map to one or more DWAs, these DWA-level labels were first combined into task-level scores, then aggregated to the occupation-level. When tasks lacked associated DWAs or when the DWA-derived scores were ambiguous, additional direct task-level labels were collected. GPT-4 rated task/occupation pairs directly, bypassing the DWA-to-task step. Both were aggregated to produce multiple occupation-level exposure scores (α, β, ζ), with core tasks weighted twice as heavily as supplemental tasks.
Hampole et al. (2025)	Task	Exposure	Created O*NET occupation-level scores by calculating the mean AI exposure probability across an occupation's tasks (weighted by O*NET importance) and the concentration (variance) of that exposure.
Handa et al. (2025)	Task	Automation and Augmentation	Measured the "depth" of AI use by calculating the fraction of tasks within each O*NET occupation that appeared in the interaction data (i.e., Claude conversations).
Kochhar (2023)	Task	Exposure	Ranked O*NET occupations by the relative importance of the exposure tiers (low, medium, high). The final "most exposed" label was assigned to occupations in the top 25% of the ranking based on high-exposure GWA importance scores.
OECD (2025)**	Task	Exposure	See Table note
Shao et al. (2025)	Task	Automation and Augmentation	Aggregated O*NET task-level Human Agency Scale (HAS) ratings to the occupation-level by examining the full distribution of both worker-desired and AI expert-assessed HAS levels across tasks within each occupation. A "dominant" HAS level (the most common category) is used as a summary metric, but the analysis also compares the shapes of the two distributions to quantify gaps between worker and expert views.
Tomlinson et al. (2025)	Task	Automation and Augmentation	Computed an AI applicability score for O*NET occupations by taking the weighted fraction of their O*NET IWAs (those appearing > 0.05% of the data) that appeared in Microsoft Bing Copilot conversations while adjusting for success rates and the scope of impact.

Table 3. (Continued)

Reference	Target	Impact Type	Method Summary
Webb et al. (2019)	Task	Automation	Calculated O*NET occupation-level exposure as a weighted average of task-level scores where task exposure is derived from text overlap between O*NET tasks and AI-related patents, where weights reflect task importance, relevance, and frequency.
Chopra et al. (2025)	KSAO	Automation	Weights each skill for a given O*NET occupation by its relative importance, automatability, and prevalence. This produces a single AI exposure value between 0 and 100%.
Felten et al. (2021)	KSAO	Exposure	Computed O*NET occupation-level AI exposure as a weighted average of O*NET ability-level exposure scores, weighted by each ability's importance and prevalence within that occupation, and scaled by the total importance-prevalence weight across all abilities in the occupation to capture relative exposure.
Fossen & Sorgner (2019)***	KSAO and Task	Multiple***	Assigned pre-existing occupation-level AI impact indices (see Frey & Osborne [2017], Felten et al., [2018], and Brynjolfsson et al., [2018]) to individuals based on their occupation and linked these to the Census's Current Population Survey (CPS) panel data to estimate effects on employment transitions and wage growth.
PwC (2025)	KSAO	Automation and Augmentation	Filtered to the top half of AI-Exposed Occupations (AIOE index values > 0.5), then classified them as augmented or automated based on whether their IMF complementarity score was above or below 0.5. For the skills-based analysis, occupation-level skill change is computed from job postings, but results are reported only as aggregates across occupation groups (e.g., AI exposure quartiles or augmentation categories), not at the level of specific occupations.
Ahmadi et al. (2024)	Vacancy	Exposure	Mapped vacancy titles to the closest O*NET titles using cosine similarity and fuzzy string matching, with remaining unmatched postings classified using an ML model. Authors then averaged topic probabilities by O*NET job family to determine the importance of each AI skill set within each family.
Alekseeva et al. (2021)	Vacancy	Exposure	Calculated an occupation-level "AI Share" as the proportion of job postings within each Standard Occupational Classification (SOC) occupation that require at least one AI-related skill, defined using a predefined list of AI skills, relative to all postings in that occupation.

Table 3. (Continued)

Reference	Target	Impact Type	Method Summary
Mäkelä & Stephany (2025)	Vacancy	Automation and Augmentation	Aggregated postings to the occupation-year level and used the log count of AI-related postings, defined via AI skill requirements, as a measure of AI adoption. They then regressed log-transformed counts of non-AI postings demanding complementary or substitutable skills on this AI adoption measure. Separately, posting-level regressions examined whether AI role status was associated with demand for complementary or substitutable skills.
Sigelman et al. (2026)	Vacancy	Automation and Augmentation	It is implied that the authors assigned AI replacement and enhancement scores to each skill within an occupation and then aggregated those skill-level exposures in an unspecified manner to yield occupation-level automation and augmentation scores.

Note. *Argarwal et al. (2025) was one of the few studies we examined that did not explicitly explore AI impact at the occupation-level. **While linkages to occupations have not yet been made, OECD proposed the "Catch-Up Index", which is computed at the capability level as the gap between AI and human requirements, and illustrated how occupation-level AI impact can be derived by taking an importance-weighted average of these scores across capabilities relevant to the job. ***Authors used multiple indices, which have different AI targets and impact types (i.e., Frey & Osborne's [2017] index, Felten et al.'s [2018] index, and Brynjolfsson et al.'s [2018] index). Target refers to the work element that the study investigated regarding AI impact. Impact Type refers to whether the researcher's AI impact metric(s) focus primarily on AI impact in a general sense (e.g., exposure) or if they made a more explicit distinction between AI automation and AI augmentation (or explored both of these).

Drawing Inferences About AI's Impact on Occupations via Job Tasks

One of the most common ways to estimate AI impact is by computing scores at the task-level (or using higher level task-like statements such as O*NET's GWAs, IWAs, or DWAs; Hansen et al., 2014) and then aggregating them up to the occupation-level (Figure 1, Path 3B). A notable example is the study by Eloundou et al. (2024). In this study, researchers developed a rubric to estimate AI "exposure", with exposure being defined as "whether access to a GPT or GPT-powered system would reduce the time required for a human to perform a specific O*NET DWA or complete a task by at least 50 percent" (page 7). Both human raters and an LLM (i.e., GPT-4) then reviewed and classified DWAs and O*NET tasks using this rubric (though humans only classified a subset of tasks). This information was then rolled up to the occupation-level, which yielded three exposure scores for each occupation: α (the strictest measure which only counts tasks where ChatGPT could help right now with no extra tools needed), β (the middle estimate which counts direct ChatGPT tasks but only giving half credit to tasks that would require additional tools to be built first), and ζ (the most generous measure in that it counts both direct ChatGPT tasks and tasks requiring additional tools). Note that greater weight was given to core tasks than to supplemental tasks (as defined in O*NET; see https://www.onetonline.org/help/online/scales) for each occupation. Collectively, these scores represent how many of an occupation's tasks could be accelerated by exposure to AI tools.

As another example, in a study conducted by Kochhar (2023), human raters classified all 41 O*NET GWAs into three categories of AI exposure (low, medium, or high) based on the question, "What is the likelihood that a work activity may be substituted for or complemented by AI at this time?" (page 9). This yielded 16 high-exposure, 16 medium-exposure, and 9 low-exposure GWAs. For each occupation, the O*NET importance rating across the GWAs within each exposure category was calculated, and the relative importance of the high/low exposure GWAs was computed and compared to all GWAs combined. Occupations were then ranked by these relative importance scores. The top 25% of occupations ranked by prominence of high-exposure GWAs were labeled as the most exposed to AI, while the top 25% ranked by prominence of low-exposure GWAs were labeled as the least exposed to AI.

In a different study, Webb (2019) extracted verb-noun pairs (e.g., “recognize image”) within AI patents and examined the semantic overlap between these extractions and O*NET task statements. A task’s AI exposure score was then computed based on how frequently the verb-noun pairs extracted from the O*NET description appeared in the AI patents. This was then aggregated to the occupation-level using weighted averages based on the task importance and task frequency ratings provided in O*NET.

Studies that focus on AI usage have also involved aggregated task-level AI impact estimates at the occupation-level. For example, Handa et al. (2025) rolled up millions of real-world AI conversations (i.e., via Claude) to the occupation-level by calculating the “depth” of usage, defined as the fraction of unique O*NET tasks within an occupation that appear in the Claude conversation data, and then categorized occupations by the share of their tasks where AI use was observed. Likewise, Agarwal et al. (2025) mapped Claude interactions to O*NET tasks but used multivariate techniques, such as principal component analysis and k-means clustering, to identify latent “task archetypes” (i.e., procedural and analytical work, dynamic problem solving, and standardized operational tasks).

Moreover, Appel et al. (2026) and Tomlinson et al. (2025) utilized different methods that incorporated additional information. For instance, Appel et al. (2026) explored the effective AI coverage for occupations by calculating the weighted sum of task success rates derived from their “economic primitives” (i.e., foundational measures of how Claude is used). Each task’s weight is determined by its share of human hours adjusted by O*NET importance and frequency scores. Additionally, Tomlinson et al. (2025) computed an AI applicability score for occupations by identifying the weighted fraction of O*NET IWAs that appear non-trivially in Copilot conversations. This score is computed by aggregating the task completion rate and the impact scope (i.e., the proportion of IWAs in which the AI assists), weighted by O*NET’s importance and relevance metrics for each occupation-IWA pair.

Drawing Inferences About AI’s Impact on Occupations via Job-Relevant KSAOs

Other research has used the AI impact estimates at the KSAO level to derive an occupation-level AI impact score (Figure 1, Path 4B). An example of this is the study by Felten et al. (2021) on the development of the Artificial Intelligence Occupational Exposure (AIOE) Index. As noted earlier, for this study, AI was operationalized using the Electronic Frontier Foundation’s (EFF) 10 AI applications/definitions (e.g., language modeling, translation, image recognition). From here, the EFF applications were linked to 52 O*NET abilities via crowdsourced ratings, with a binary (i.e., 1/0) scale indicating whether an EFF application and an O*NET ability were linked. Once AI exposure estimates were derived for the abilities, for each occupation, researchers calculated an AIOE Index score by weighting each ability’s AI exposure by its prevalence and importance within that occupation, then summing across all abilities and scaling by the total number of abilities required within that occupation.

Moreover, PwC (2025) extended this study by Felten et al. (2021) by focusing on skills and integrating the AIOE Index with the IMF Complementarity Index (CI; Pizzinelli et al., 2023) to more explicitly distinguish between AI automation and AI augmentation. Specifically, occupations were first filtered to include only those with an AIOE score greater than 0.5, representing the upper half of AI-exposed occupations, and then further classified using the IMF CI, which was rescaled to fall between 0 and 1. Occupations scoring above 0.5 on the CI were classified as augmented, indicating that AI enhances rather than replaces human work, while those scoring below 0.5 were classified as automated, indicating that AI serves as a substitute. Separately, PwC (2025) also used Lightcast's AI-skills taxonomy to identify postings requiring AI skills and observed changes over time in the prevalence of AI-related job postings and associated skill demand. Additionally, the "Iceberg Index" developed by Chopra et al. (2025) measured AI impact across occupations using a skills-centered approach. This index relies on three dimensions: (1) the skills required by each O*NET occupation, (2) the extent to which these skills can be automated by AI based on cataloged AI tools (over 13,000 AI tools were cataloged in this research), and (3) the economic value of the work (i.e., wages and employment). For each occupation, the index weighs each skill by its relative importance, automation score, and prevalence to produce a single exposure value between 0-100% that quantifies the "wage value of skills that AI systems can perform within each occupation, revealing where human and AI capabilities overlap" (page 6).

Drawing Inferences About AI's Impact on Occupations via Job Vacancies

Finally, studies have also directly estimated the impact of AI at the occupation-level using job vacancy data (Figure 1, Path 5A). For example, Ahmadi et al. (2024) examined the demand for ChatGPT-related skills by analyzing 1,128 job advertisements from Indeed, LinkedIn, and ZipRecruiter. Researchers mapped job postings to O*NET occupation titles using a two-step process with Term Frequency-Inverse Document Frequency (TF-IDF) cosine similarity matching, fuzzy string matching, and an ML classification model. They then used topic modeling on ChatGPT-specific portions of job descriptions to identify five distinct skill sets (e.g., creative content generation), which were then mapped back to O*NET job families to determine the relative centrality of each skill set within each occupation group. Likewise, Mäkelä and Stephany (2025) examine the impact of AI (both automation and augmentation) using approximately 12 million U.S. online job vacancies. The authors identified "AI roles", which were defined as roles requiring at least one of 115 specific AI skills across various categories (e.g., ML, NLP, computer vision, and image processing). The authors then aggregated job posting data to the occupation-level by grouping individual job postings according to their Standard Occupational Classification (SOC) codes. From here, they calculated the share of postings within each occupation-year combination that required AI-augmentative or AI-automated skills.

Summary of Methodological Approaches to AI Impact

In sum, most studies on AI impact can be mapped to a specific sequence of paths within Figure 1. In Table 4 below, we provide an *AI Impact Method Profile* for each study reviewed above, denoting the specific paths from Figure 1 used by each study to derive AI impact estimates, with most culminating in occupation-level estimates. As noted above, many studies start with a set of AI capabilities and first draw inferences about AI impact at the task-level, before aggregating to the occupation-level (e.g., Brynjolfsson et al., 2018; Eloundou et al., 2024; Hampole et al., 2025, Kochhar, 2023; OECD, 2025; Shao et al., 2025; Webb et al., 2019). Alternatively, some studies examine AI usage and use this as a basis for inferring AI's impact on job tasks and, in turn, occupations (e.g., Agarwal et al., 2025; Appel et al., 2026; Handa et al., 2025; Tomlinson et al., 2025). Other studies start with a set of AI capabilities and first draw inferences about AI impact at the KSAO level before aggregating to the occupation-level (e.g., Chopra et al., 2025; Felten et al., 2021; Fossen & Sorgner, 2019; PwC, 2025). Lastly, a final set of studies center on the use of job vacancy data (e.g., job advertisements or job postings) and uses that as a basis for computing AI impact scores at the occupational level (e.g., Ahmadi et al., 2024; Alekseeva et al., 2021; Mäkelä & Stephany, 2025; Sigelman et al., 2026). Of course, each of these studies vary in their specifics (e.g., there specific operationalizing of AI, how they defined/scaled AI impact, how they actually linked AI capes/usage to specific elements of work; see Tables 1 through 3), but Figure 1 and Table 4 help illustrate that these studies can be meaningfully clustered from a methods perspective in terms of how they went from an initial operationalization of AI to draw inferences about AI's impact on work.

Table 4. AI Impact Method Profiles for Reviewed Studies Based on Figure 1

Reference	AI Impact Method Profile (based on Figure 1 Boxes/Paths)													
	1	2	1A	1B	1C	2A	3	3A	3B	4	4A	4B	5	5A
AI Capes → Tasks → Occupations														
Brynjolfsson et al. (2018)	Y	--	Y	--	--	--	Y	Y	Y	--	--	--	--	--
Eloundou et al. (2024)	Y	--	Y	--	--	--	Y	Y	Y	--	--	--	--	--
Hampole et al. (2025)	Y	--	Y	--	--	--	Y	Y	Y	--	--	--	--	--
Kochhar (2023)	Y	--	Y	--	--	--	Y	Y	Y	--	--	--	--	--
OECD (2025)*	Y	--	--	--	--	--	--	--	--	--	--	--	--	--
Shao et al. (2025)	Y	--	Y	--	--	--	Y	Y	Y	--	--	--	--	--
Webb et al. (2019)	Y	--	Y	--	--	--	Y	Y	Y	--	--	--	--	--
AI Usage → Tasks → Occupations														
Agarwal et al. (2025)**	--	Y	--	--	--	Y	Y	Y	--	--	--	--	--	--
Appel et al. (2026)	--	Y	--	--	--	Y	Y	Y	Y	--	--	--	--	--
Handa et al. (2025)	--	Y	--	--	--	Y	Y	Y	Y	--	--	--	--	--
Tomlinson et al. (2025)	--	Y	--	--	--	Y	Y	Y	Y	--	--	--	--	--
AI Capes → KSAOs → Occupations														
Chopra et al. (2025)	Y	--	--	Y	--	--	--	--	--	Y	Y	Y	--	--
Felten et al. (2021)	Y	--	--	Y	--	--	--	--	--	Y	Y	Y	--	--
Fossen & Sorgner (2019)***	Y	--	Y	Y	--	--	Y	Y	Y	Y	Y	Y	--	--
PwC (2025)****	Y	--	--	Y	Y	--	--	--	--	Y	Y	Y	Y	Y
AI Capes → Vacancies → Occupations														
Ahmadi et al. (2024)	Y	--	--	--	Y	--	--	--	--	--	--	--	Y	Y
Alekseeva et al. (2021)	Y	--	--	--	Y	--	--	--	--	--	--	--	Y	Y
Mäkelä & Stephany (2025)	Y	--	--	--	Y	--	--	--	--	--	--	--	Y	Y
Sigelman et al. (2026)	Y	--	--	--	Y	--	--	--	--	--	--	--	Y	Y

Note. Column numbers correspond to the boxes/paths specified in Figure 1. 1 = AI Capabilities. 1A = AI Capabilities → Task Statements. 1B AI Capabilities → KSAOs. 1C = AI Capabilities → Vacancy-Level Descriptions. 2 = AI Usage. 2A = AI Usage → Task Statements. 3 = Task Statements. 3A = Task Statements → Task-Level AI Impact Score. 3B = Task-Level AI Impact Score → Occupation-level AI Impact Score. 4 = KSAOs. 4A = KSAOs → KSAO-Level AI Impact Score. 4B = KSAO-Level AI Impact Score → Occupation-level AI Impact Score. 5 = Job Vacancy Elements. 5A = Job Vacancy Elements → Occupation-level AI Impact Score.

*While OECD provided an initial illustration of how to link tasks and occupations to their AI capability indicators, this has not yet been done at scale (to our knowledge). **Argarwal et al. (2025) was one of the few studies we examined that did not explicitly explore AI impact at the occupation-level. ***Indices that assess different types of AI impact (e.g., both tasks and KSAOs) were used. **** PwC (2025) employed multiple metrics to classify whether jobs could be automated or augmented by AI and also explored demand for AI skills via job descriptions.

Recommendations for Developing a Suite of AI Impact Indices for O*NET

In this section, we detail our recommendations for developing a suite of AI impact indices for O*NET. As noted in the introduction, we begin by discussing key considerations and observations about past AI impact research that influenced our recommendations. Next, we introduce each index and how it could be used in the context of O*NET.

Key Considerations and Observations About Past AI Impact Research

Upon reviewing the rapidly evolving literature on AI's impact on work, we identified a set of considerations that helped shape our recommendations for maintaining AI impact data for O*NET. We think of these as fundamental design considerations that influenced our recommendations. Our main considerations are detailed in the sections that follow.

Embracing a Job Analytic Perspective

One consistent trend we observed in the AI impact approaches to date is that there is very little cross-fertilization with literature on (a) how jobs have traditionally been analyzed, (b) modern perspectives on job performance (e.g., Borman & Motowidlo, 1997; Brannick et al., 2007; Campbell & Wiernik, 2015; Dorsey et al., 2017; Wilson et al., 2012). We view this as a limitation of current approaches that we aim to address through our recommendations.

For example, the majority of AI impact work to date has adopted a very "task-centric" perspective on job performance (e.g., Brynjolfsson et al., 2018; Eloundou et al., 2024; Hampole et al., 2025, Kochhar, 2023; OECD, 2025; Shao et al., 2025; Webb et al., 2019). In contrast, modern theories and research on job performance have recognized that job performance reflects far more than "task performance" and encompasses other performance behaviors that are critical to organizational functioning (e.g., contextual performance; Borman & Motowidlo, 1993, 1997; Campbell & Wiernik, 2015). Interestingly, from an AI impact perspective, we anticipate AI impact (in general) to be far less for contextual performance behaviors and for the performance of job tasks, which suggests the ability of AI to automate or augment performance of a job may have been overstated by the current body of AI research due to its ignoring non-task, yet important elements of job performance identified to date.

Another example is that there are well-established models of job performance that recognize three proximal determinants: Declarative Knowledge (DK), Procedural Knowledge and Skill (PKS), and motivation (i.e., choice to direct and persist with effort) (e.g. Campbell et al., 1993; Motowidlo & Borman, 1997; Schmidt, 2014). When attempting to understand AI's impact on work, indices typically adopt either task performance or more Knowledge-Skill-Ability (KSA)-centric perspectives, but within a study, those perspectives are rarely directly compared or their differences elaborated upon with respect to their implications for AI's impact on work. For example, if one had an index focused on assessing whether AI had the capability to exhibit the requisite DK or PKS to perform a given job-critical task, and another focused on assessing whether AI could automate or augment the performance of that task that led to different conclusions about AI impact, what would it mean? Would the differences be sensible based on accumulated scientific knowledge on job performance and its determinants? Having multiple indices that approach the question of AI impact through the lens of modern models of job performance and its determinants will provide the O*NET Program with a foundation for offering such insights.

Adopting Micro and Macro Perspectives

Related to embracing a job-analytic perspective, there are different ways one may approach AI's impact, whether from a bottom-up, micro perspective or a top-down, macro perspective. For example, one can frame the issue from more atomic or micro assessing AI impact by linkages to the elements of a job (e.g., individual tasks, individual required knowledges or skills), or from a more macro perspective by assessing AI impact linkages to a job as a whole (e.g., impact on a job defined in terms of a constellation or set of tasks and behaviors). The latter perspective recognizes performance of a job may not simply be a "sum of its parts" but rather attempts to capture interdependencies among tasks and behaviors that serve to define performance on a job that may be lost by approaches that involve simple aggregation of AI impact on its elemental parts (e.g., aggregation of AI impact on individual job tasks). This is something that, to date, has arguably been under-examined in the AI impact literature, as most studies have adopted a more bottom-up, micro-level approach that first draws inferences about work elements and then aggregates them.

Quality of Job/Occupation-Information Used as Inputs

Another consideration is the quality of the inputs used to infer what work on a job entails. One major differentiator of existing AI impact approaches is the source of information used to help determine what work in an occupation or job involves. For example, much research has used well-vetted and incumbent-sourced data on key occupational tasks and/or KSAs from gold-standard sources such as the Department of Labor's O*NET System (e.g., Brynjolfsson et al., 2018; Eloundou et al., 2024; Felten et al., 2021; Handa et al., 2025; Shao et al., 2025; Tomlinson et al, 2025), which have long-standing data collection programs and quality control mechanisms behind such data in place (National Center for O*NET Development, 2026b).

In contrast, other studies have inferred what a job requires from online vacancy announcements (e.g., Ahmadi et al., 2024; Alekseeva et al., 2021; Mäkelä & Stephany, 2025; Siegelman et al., 2026). It is important to note that, unlike O*NET data, there are no common quality standards or requirements for the depth of information about a job that appears in vacancy announcements across employers. In many ways, one can argue that job vacancies are more of a "marketing" device for employers to attract individuals to apply for a job, rather than being rigorous, systematic descriptions of the key job tasks and the critical KSAs needed to form that job that are grounded in a systematic job analysis that meets professional standards (Hurtz & Wright, 2012). Indeed, little literature has systematically examined the quality of job vacancy data from the perspective of what constitutes a robust job description (Carnevale et al., 2014; Hurtz & Wright, 2012; Lancaster et al., 2019). Surprisingly, a search of the literature revealed no clear treatment of this issue over the past 7 years, despite increasing interest in leveraging job vacancies to infer job requirements and skills.

To help illustrate this deficiency, a simple experiment here would be to obtain a large number of vacancy announcements for a given class of jobs (e.g., those that fall within a given O*NET-SOC or SOC occupation), and use an LLM to tag them for the presence/absence of key literature-based features of a robust job description (Hurtz & Wright, 2012). Of course, even if those features were present, this would not reveal whether that information was grounded in a job analysis, but it would indicate general areas of deficiency, which would be useful for developing clear, research-informed inclusion/exclusion rules for use of job vacancy data in AI impact efforts.

Of course, a counterargument here is that job vacancy data provides insight into within-occupation heterogeneity that is masked by occupation-level data (e.g., insights into different types of positions or jobs within an occupation, and the differences in key tasks and KSAs they require), and can provide more granularity than an occupation-level source such as O*NET. Furthermore, vacancies offer a more real-time signal of tasks, skills, and demands from organizations, relative to the majority of data currently included in O*NET (Software Skills and Jobs Titles are an important exception), which tends to be more static and changes less often, while potentially offering a high volume of job-related information that is more granular and constantly refreshed as new jobs are posted. While using vacancies as sources of job-related information is a positive feature, it does not necessarily mean the information *accurately or comprehensively reflects the tasks or KSAO requirements for the job in question*. Given the observations above, our recommendations aim to consider job vacancy data as only one potential source for inferring AI's impact on work, and to be very transparent about strengths and limitations. As we note, several of our proposed indices do involve the use of vacancies; however, our approach to using vacancies for this purpose would first ensure that the vacancies that serve as input possess an acceptable level of quality, as reflected in their inclusion of literature-based features of a robust job description (Hurtz & Wright, 2012).

Limitations of Human Experts to Make AI Impact Judgments

As noted in our review, various studies are reliant on subjective human judgment to index AI impact (e.g., Eloundou et al., 2024; Felten, 2021; Kochhar, 2023; OECD, 2025), and there are potential limitations of doing so that may create a basic validity issue for the ratings gathered. The dilemma here is that the expertise to make a meaningful rating often does not reside within a single human SME. For example, a given human SME may bring expertise in a given job (e.g., an incumbent or supervisor in a given job) yet may have insufficient understanding of the capabilities of AI to make a valid rating of AI's impact on that job. Conversely, another human SME may bring expertise in AI capabilities but lack sufficient understanding of the job's nuances, job performance requirements, and the suite of KSAOs needed to function effectively in that job, to make a valid rating of AI's impact for that job. Though ensuring a mix of expertise among human SMEs tasked with making these judgments can help, it does not necessarily offset the issue that the ratings from each SME are based on what may be very partial (deficient) understanding of factors needed to make valid judgments (i.e., averaging ratings across SMEs can help reduce idiosyncratic errors among SMEs to develop final aggregate scores, but not systematic deficiencies in their ratings). Of course, the other issue with human SME ratings is that they can be very time-, resource-, and cost-intensive to collect and, as such, very difficult to use as a basis for maintaining AI impact ratings for the 800+ data-level occupations that may exist in O*NET at any point in time. In contrast, it is possible for both job-specific and AI capability "knowledge" to exist with the knowledge base upon which an LLM is prompted, or reflected in the context an LLM is given if tasked to make ratings of AI's impact on a given job, and also far more feasible to scale to maintain data for 800+ occupations relative to human-based ratings.

Limitations of Usage-Based AI Impact Metrics

As noted in the review section of this report, multiple past AI impact studies have framed the question of AI impact from the perspective of actual usage data from applications such as Microsoft Bing Copilot or Anthropic's Claude platform (e.g., Appel et al., 2026; Handa et al., 2025; Tomlinson et al., 2025). From one perspective, these studies offer a critical perspective on the AI impact question by examining actual AI-related usage behavior (rather than human- or LLM-based "theoretical" judgments of AI's impact). While this such work clearly fills a void it is limiting in that it the applications examined reflect relatively narrow ways one may interface with AI and how AI can be leveraged at work (e.g., example, contrast these chat base applications with the range of AI capabilities formulated by OECD, 2025) raising concerns about the extent to which findings generalize to cover the full range of AI usage that go well beyond chat-focused applications.

Additional Considerations

The areas we highlighted can be viewed as high-level observations we made regarding past AI impact work that directly relate to the recommendations we make to the O*NET Program later in this section. In addition to those observations, we considered the following questions as we formulated our recommendations for a suite of AI impact indices for the Program.

- **Does the AI impact index lend itself to a clear evaluation strategy?** Another key consideration is that we aimed to develop indices that lend themselves to a clear evaluation strategy (e.g., convergence with other similar indices or expert ratings).
- **How easy and efficient would it be to refresh the AI impact index over time for hundreds of O*NET occupations?** Another important consideration is the ease and efficiency with which any proposed method can be implemented to refresh AI impact ratings over time (e.g., in the face of changes to occupations, job tasks, knowledge, or skills, or in light of the introduction of new occupations).
- **Is the suite of proposed AI impact indices able to support multiple potential units of analysis and, in turn, use cases within the O*NET ecosystem?** This consideration involves allowing flexibility to make inferences about AI impact at different unit analyses that the O*NET Program may want to convey within the O*NET ecosystem (e.g., presenting AI impact data not only for an occupation in general, but also for individual tasks, knowledge, or skills [independently] within an occupation). For example, beyond reporting AI's impact on occupations, as we discuss later, one possibility is to annotate existing O*NET skills (software, essential, and transferable) with indicators of which have a nexus to or may be impacted by AI.
- **How free is the AI impact index from external dependencies?** This consideration involves how free or dependent a proposed method is on maintaining structured data from sources outside the O*NET Program's or DOL's control. For example, if the method relies on data currently provided by an external entity and there are no published plans to document how that data will be maintained or updated in the future, it creates risk for the Program to rely on that source for updating O*NET data.
- **To what extent does the AI impact index leverage existing resources and past work?** This consideration involves leveraging existing code and resources where they already exist. For example, if there are open-source, validated (for the purposes at hand), and publicly available prompts, code, or processes for linking AI capabilities to tasks, we would look to leverage such resources rather than developing something new from the ground up.

Proposed AI Impact Indexes

In light of the observations and considerations above, and our review of methods for indexing AI impact to date, we propose a suite of 16 AI impact indices for O*NET to consider exploring that vary along three main dimensions:

- **Target Focus** (job performance *determinant* vs. *component* focused). We organize our indices by the type of element within a literature-based job performance model, focusing on addressing the question of AI impact. Specifically, our indices differ in terms of whether they target *determinants* of performance (i.e., job-required DK and PKS), *components* of job performance (i.e., task vs. contextual performance), or overall job performance. We comment on this further below.
- **Unit of Analysis** (job element vs. vacancy-level vs. occupation-level). A secondary dimension along with which indices can be organized is the unit of analysis for which they produce AI impact ratings. Specifically, the proposed indices differ in whether the unit of analysis is a specific knowledge, skill, task, or job behavior (i.e., job element-level units), a job vacancy (i.e., a specific job or posting within an occupation), or an occupation (specifically defined by O*NET-SOC).
- **Impact Type** (augmentation vs. automation). A final dimension along which most of our indices differ is in terms of AI impact type, namely automation vs. augmentation, recognizing that conflating these two distinct phenomena risks obscuring meaningfully different implications for workers and occupations. Specifically:
    - *AI augmentation* refers to using AI to assist people with work they already perform, improving speed, accuracy, and clarity while keeping humans fully responsible for decisions (i.e., a "co-pilot" model).
    - *AI automation* refers to AI systems that act without human approval, interpreting information, choosing the next step, and executing it independently, effectively replacing human labor on those tasks.

  Given these important distinctions, the suite is designed so that each automation-focused index has a parallel augmentation-focused counterpart, allowing the O*NET Program to present a more complete and nuanced view of AI's potential impact on work.

As noted above, the main organizing dimension for our proposed indices is which element of contemporary models of job performance and its determinants they target. Figure 2 provides a conceptual mapping between the proposed AI impact indices and a contemporary model of job performance and its determinants (e.g., Borman & Motowidlo, 1997; Campbell et al., 1993), and Table 5 provides a high-level description of each AI impact index. Some indices focus on AI's impact on proximal determinants of performance (i.e., job-required DK and PKS), whereas others focus on AI's impact on major components of performance (i.e., task and contextual performance) or adopt a more general focus on overall job performance. In this context, declarative knowledge refers to knowledge of facts, rules, principles and procedures needed to perform effectively (i.e., knowing what to do), whereas procedural knowledge and skill is the "capability attained when declarative knowledge has been successfully combined with knowing how (i.e., knowing how to do it) and being able to perform at task" (adapted from Kanfer & Ackerman, 1990; McCloy et al., 1994).

Note that we do not propose developing AI impact indices corresponding to the "motivation" proximal determinant of performance, as that represents human volitional choice to direct and persist with effort, which is a factor when modeling one's typical performance, but is less relevant when modeling maximal human job performance (e.g., McCloy et al., 1994). We assume that, when it comes to AI, motivation (as traditionally defined) is not as clearly relevant for AI as it is for humans in determining job performance (or, at best, would manifest very differently for AI; see the *Potential Future Considerations* section towards the end of this report).

Additionally, we also do not attempt to provide AI impact indices for distal performance determinants (i.e., individual difference constructs and training, experience, and education that facilitate or lead to the acquisition of DK or PKS or may affect one's motivation to perform a given job, such as one's specific interests and values). The reason for the latter is twofold. First, per Figure 2, the model of performance determinants implies that distal determinants affect job performance only (or largely) via proximal determinants. Thus, to the extent this model reflects reality, if we address the question of AI's impact on DK and PKS, then considering AI's impact on more distal precursors of DK and PKS will not add value to understanding AI's impact on job performance beyond the DK and PKS required for the job. The second reason is more practical: given their indirect relation to job performance, we anticipated diminishing returns from adding more indices to what was already a large, proposed set.

Figure 2. Mapping of Proposed O*NET AI Impact Indices onto a Contemporary Model of Job Performance and Its Determinants

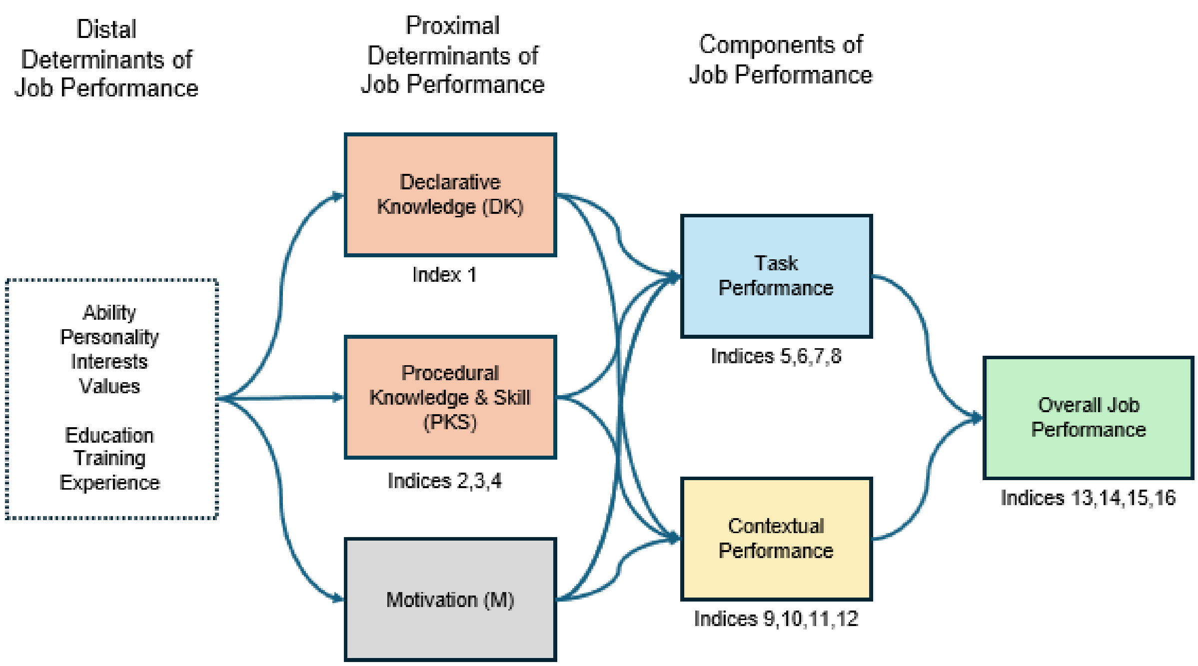


Table 5 provides an overview of the full suite of proposed job-analysis-centric AI impact indices. As shown in Table 5, the proposed suite comprises 16 distinct indices, organized by the job performance model element in Figure 2 that each targets. Also included in Table 5 are initial thoughts on how each index may be used within the O*NET ecosystem and the potential benefits it could provide to O*NET users.

Though not the focus of this report, our current thinking is that the approach for generating all these indices would involve prompting multiple LLMs to provide structured ratings, using different types of inputs that are either O*NET-derived or sourced from job vacancies, and then aggregating those ratings across models. Using multiple LLMs rather than a single model would guard against idiosyncratic biases or limitations of any single model, and the variance observed across models can itself be informative for assigning uncertainty estimates to ratings for each index. For each index, LLM prompts would be carefully constructed to include: (a) an operationalization of AI, (b) the relevant occupation/job/job-element input content (e.g., a task statement, a job vacancy, a skill definition and its level anchors, an occupation description, or a contextual behavior); (c) a precise definition of the AI impact construct being rated (e.g., automation or augmentation); and (d) a structured rating scale with clearly specified anchors (where appropriate). Note that the operationalization of AI may not simply be limited to information specified in the prompt itself but could reference an external context that the prompt is asked to consider (e.g., an LLM-generated white paper crafted on the current state of AI applications to work at the time AI impact ratings are to be refreshed). Additionally, note that the distinction between automation and augmentation would be made explicit in prompts that target job performance or its components. Augmentation refers to AI that assists and enhances human performance while keeping humans responsible for decisions, while automation refers to AI systems that act without human approval, interpreting information, choosing next steps, and executing them independently. In the sections that follow, we describe each cluster of indices, focusing our overview on their conceptual rationale and connection to the broader performance model in Figure 2, and largely leaving summaries of their potential use and benefit to the information provided in Table 5.

Table 5. Overview of Proposed AI Impact Indices

AI Impact Indices	Unit of Analysis	Example Potential Use in O*NET	Potential Benefit to O*NET Users
Performance Determinant Focused Indices			
Index 1: Knowledge-level index focused on evaluating whether the current state of AI technology can exhibit a specified level of job-relevant knowledge	Job Element - Knowledge	Annotate important knowledges in an occupation's O*NET Online profile with an indicator/rating of the potential for AI to exhibit the required level of knowledge for that occupation Publish indicators/ratings along with knowledge importance and level ratings for an occupation to the O*NET database Aggregate indicators/ratings to occupation-level to provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database	When reviewing an occupation, users can get a sense of what knowledges within an occupation may be most/least impacted by AI Enriches the O*NET database with additional knowledge-level data and an accompanying aggregate occupation-level metric to facilitate users' decision making, applications, and research
Index 2: Skill-level index focused on evaluating whether the current state of AI technologies can exhibit a specified level of job-relevant essential or transferable skill	Job Element – Essential and Transferable Skill (O*NET)	Important, essential, and transferable skills for an occupation in O*NET Online could be tagged with an indicator or rating of the potential for AI to exhibit the required level of skill for performing the occupation in question Publish indicators/ratings along with skill importance and level ratings for an occupation to the O*NET database Aggregate indicators/ratings to occupation-level to provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database	When reviewing an occupation, users can get a sense of what essential and transferable skills within an occupation may be most/least impacted by AI Enriches the O*NET database with additional skill-level data and an accompanying aggregate occupation-level metric to facilitate users' decision making, applications, and research. Facilitates understanding of AI-skill connections for those interested in skills-based hiring
Index 3: Skill-level index focused on evaluating whether a given O*NET software skill is AI-related	Job Element – Software Skill (O*NET)	Software skills for an occupation in O*NET Online could be tagged with an indicator of whether the software skill is AI-related Publish AI indicators for each software skill in an occupation's software skill list to the O*NET database Aggregate indicators at the occupation level to provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database	When reviewing an occupation, users can get a sense of what software skills within an occupation have a nexus to AI Enriches the O*NET database with additional skill-level data and an accompanying aggregate occupation-level metric to facilitate users' decision making, applications, and research. Facilitates understanding of AI-skill connections for those interested in skills-based hiring

AI Impact Indices	Unit of Analysis	Example Potential Use in O*NET	Potential Benefit to O*NET Users
Index 4: Skill-level index focused on evaluating whether a given vacancy-sourced skill is AI-related	Job Element – Skill (Vacancy)	Granular, vacancy-based skills for an occupation could be tagged with an indicator of whether the software skill is AI-related Publish AI indicators for each vacancy-sourced skill in a new vacancy-sourced skill list for occupations in the O*NET database Aggregate indicators at the occupation-level to provide one element of a new occupation-level AI Impact section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database	When reviewing an occupation, users can get a sense of what vacancy-sourced skills within an occupation have a nexus to AI Enriches the O*NET database with additional, more granular vacancy-sourced skill-level data and an accompanying aggregate occupation-level metric to facilitate users' decision making, applications, and research. Facilitates understanding of AI-skill connections for those interested in skills-based hiring
Task Performance Focused Indices			
Index 5: Job task-level index focused on evaluating whether the current state of AI technology can AUTOMATE performance of a job task	Job Element - Task	Annotate important tasks in an occupation's O*NET Online profile with an indicator/rating of the potential for AI to automate performance of the task Publish indicators/ratings along with task importance and frequency ratings for an occupation to the O*NET database Aggregate indicators/ratings to occupation-level to provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database	When reviewing an occupation, users can get a sense of what tasks within an occupation may be most/least susceptible to automation by AI Enriches the O*NET database with additional task-level data and an accompanying aggregate occupation-level metric to facilitate users' decision making, applications, and research
Index 6: Job task-level index focused on evaluating whether the current state of AI technology can AUGMENT humans' performance of a job task	Job Element - Task	Annotate important tasks in an occupation's O*NET Online profile with an indicator/rating of the potential for AI to augment humans' performance of the task Publish indicators/ratings along with task importance and frequency ratings for an occupation to the O*NET database Aggregate indicators/ratings to occupation-level to provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database	When reviewing an occupation, users can get a sense of what tasks within an occupation AI may show the most/least potential for augmenting humans' performance of the task Enriches the O*NET database with additional task-level data and an accompanying aggregate occupation-level metric to facilitate users' decision making, applications, and research

AI Impact Indices	Unit of Analysis	Example Potential Use in O*NET	Potential Benefit to O*NET Users
Index 7: Occupation-level index focused on evaluating whether the current state of AI technology can AUTOMATE performance of an occupation based on its description and a collective set of job tasks	Occupation	Provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database. This element would offer a holistic estimate of the potential for AI to automate task performance in the occupation	When reviewing an occupation, users can get an overall sense of the potential for AI to automate task performance in the occupation Enriches the O*NET database with a new occupation-level metric to facilitate users' decision making, applications, and research
Index 8: Occupation-level index focused on evaluating whether the current state of AI technology can AUGMENT humans' performance of an occupation based on its description and a collective set of job tasks	Occupation	Provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database. This element would offer a holistic estimate of the potential for AI to augment humans' task performance in the occupation	When reviewing an occupation, users can get an overall sense of the potential for AI to augment humans' task performance in the occupation Enriches the O*NET database with a new occupation-level metric to facilitate users' decision making, applications, and research
Contextual Performance Focused Indices			
Index 9: Contextual behavior-level index focused on evaluating whether the current state of AI technology can AUTOMATE performance of a job-relevant contextual performance behavior	Job Element - Contextual Behavior	Provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database. This element would offer an aggregate (bottom-up) estimate of the potential for AI to automate contextual performance in the occupation Can also provide the basis for highlighting an example set of occupationally-relevant contextual behaviors (and AI's estimated impact on them) for a given occupation – serving to complement an occupation's O*NET task list.	When reviewing an occupation, users can get a sense of automation potential with respect to not only task performance (a historical focus of AI impact automation indices), but also with respect to contextual performance - a critical, yet often overlooked element of job performance Enriches the O*NET database with a new aggregate occupation-level metric to facilitate users' decision making, applications, and research
Index 10: Contextual behavior-level index focused on evaluating whether the current state of AI technology can AUGMENT humans' performance of a job-relevant contextual performance behavior	Job Element - Contextual Behavior	Provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database. This element would offer an aggregate (bottom-up) estimate of the potential for AI to augment humans' contextual performance in the occupation Can also provide the basis for highlighting an example set of occupationally-relevant contextual behaviors (and AI's estimated impact on them) for a given occupation – serving to complement an occupation's O*NET task list	When reviewing an occupation, users can get a sense of augmentation potential with respect to not only task performance (a historical focus of AI impact augmentation indices), but also with respect to contextual performance - a critical, yet often overlooked element of job performance Enriches the O*NET database with a new aggregate occupation-level metric to facilitate users' decision making, applications, and research

AI Impact Indices	Unit of Analysis	Example Potential Use in O*NET	Potential Benefit to O*NET Users
Index 11: Occupation-level index focused on evaluating whether the current state of AI technology can AUTOMATE the performance of an occupation based on its description and a collective set of contextual behaviors	Occupation	Provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database. This element would offer a holistic estimate of the potential for AI to automate contextual performance in the occupation	When reviewing an occupation, users can get a sense of automation potential with respect to not only task performance (a historical focus of AI impact automation indices), but also with respect to contextual performance - a critical, yet often overlooked element of job performance Enriches the O*NET database with a new occupation-level metric to facilitate users' decision making, applications, and research
Index 12: Occupation-level index focused on evaluating whether the current state of AI technology can AUGMENT humans' performance of an occupation based on its description and a collective set of contextual behaviors	Occupation	Provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database. This element would offer a holistic estimate of the potential for AI to augment humans' contextual performance in the occupation	When reviewing an occupation, users can get a sense of augmentation potential with respect to not only task performance (a historical focus of AI impact augmentation indices), but also with respect to contextual performance - a critical, yet often overlooked element of job performance Enriches the O*NET database with a new occupation-level metric to facilitate users' decision making, applications, and research
Overall Job Performance Focused Indices			
Index 13: Occupation-level index focused on evaluating whether the current state of AI technology can AUTOMATE performance of an occupation based on its description and a collective set of job tasks and contextual behaviors	Occupation	Provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database. This element would offer a holistic estimate of the potential for AI to automate performance of the occupation (factoring in task and contextual performance)	When reviewing an occupation, users can get an overall sense of the potential for AI to automate performance of the occupation (again, factoring in task and contextual performance) Enriches the O*NET database with a new occupation-level metric to facilitate users' decision making, applications, and research

AI Impact Indices	Unit of Analysis	Example Potential Use in O*NET	Potential Benefit to O*NET Users
Index 14: Occupation-level index focused on evaluating whether the current state of AI technology can AUGMENT humans' performance of an occupation based on its description and a collective set of job tasks and contextual behaviors	Occupation	Provide one element of a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database. This element would offer a holistic estimate of the potential for AI to augment humans' performance of work in the occupation (factoring in task and contextual performance)	When reviewing an occupation, users can get an overall sense of the potential for AI to augment humans' performance of work in the occupation (again, factoring in task and contextual performance) Enriches the O*NET database with a new occupation-level metric to facilitate users' decision making, applications, and research
Index 15: Vacancy-level index focused on evaluating whether the current state of AI technology can AUTOMATE performance of the job as described in the vacancy	Vacancy	Provide multiple elements for a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database. Vacancy-level AI impact data could be summarized and visualized at the occupation-level in several ways within an occupation's O*NET Online profile (e.g., counts/ percentage of vacancies for an occupation within the latest time period where AI was deemed as having some minimum potential to automate work; trends in such counts/percentages across multiple time periods for a given occupation)	When reviewing an occupation, users can get an overall sense of the extent to which the potential for AI to automate performance of work appears evident across job vacancies for the occupation (both based on the latest vacancies and trends in vacancies over time Enriches the O*NET database with new longitudinally focused, vacancy-driven, occupation-level metrics to facilitate users' decision making, applications, and research
Index 16: Vacancy-level index focused on evaluating whether the current state of AI technology can AUGMENT humans' performance of the job as described in the vacancy	Vacancy	Provide multiple elements for a new *Occupation-level AI Impact* section of an occupation's profile on O*NET Online, and a new AI-focused section of the O*NET Content Model and database. Vacancy-level AI impact data could be summarized and visualized at the occupation-level in several ways within an occupation's O*NET Online profile (e.g., counts/ percentage of vacancies for an occupation within the latest time period where AI was deemed as having some minimum potential to augment humans' performance of work; trends in such counts/percentages across multiple time periods for a given occupation)	When reviewing an occupation, users can get an overall sense of the extent to which the potential of AI to augment humans' performance of work appears evident across job vacancies for the occupation (both based on the latest vacancies and trends in vacancies over time) Enriches the O*NET database with new longitudinally focused, vacancy-driven, occupation-level metrics to facilitate users' decision making, applications, and research

Performance Determinant-Focused Indices (Indices 1–4)

The first four indices focus on the proximal determinants of job performance, specifically, the DK and PKS elements of the performance model in Figure 2. Rather than asking whether AI can perform a task, the first two of these indices focus on a more foundational question: Can the current state of AI technology demonstrate the level of knowledge or skill required by a given occupation? This distinction is important. An AI system might be capable of performing a task without possessing the underlying knowledge or skill in any meaningful sense, just as a sophisticated calculator can produce correct arithmetic without understanding mathematics. Conversely, an AI system may demonstrate impressive domain knowledge without being able to translate that knowledge into skilled task execution. By developing separate indices for knowledge and skill, the suite allows the O*NET Program to disentangle these contributions and offer a richer, more theoretically grounded portrait of AI's relationship to the knowledge and skill determinants of performance at work.

Index 1 is a declarative knowledge-level index focused on evaluating whether the current state of AI technology can exhibit a specified level of job-relevant knowledge as reflected by the set of O*NET Knowledge elements. This index represents a conceptual variant on the OECD “AI capabilities indicator” approach (OECD, 2025) but critically inverts its orientation. Rather than characterizing AI capabilities in terms of dimensions of AI capability, it situates the rating question directly within the O*NET Knowledge domain by using existing O*NET Knowledge element definitions, an occupationally relevant knowledge-level anchor, and existing level anchors as inputs. Similarly, *Index 2* is a parallel skill-level index, focused on evaluating whether the current state of AI technologies can exhibit a specified *level* of job-relevant essential or transferable skill as reflected by the set of those O*NET skill elements.

In addition to offering essential and transferable skills for occupations, O*NET also provides lists of software skills; unlike essential and transferable skills, these software skills are not accompanied by importance or level ratings. As such, Index 3 shifts its strategy and focuses on evaluating whether a given O*NET software skill is AI-related. Lastly, our final index in this set (Index 4), is also skill-focused, addressing more granular-level skills solely sourced from vacancies. As a set, Indices 2-4 could facilitate understanding of AI-skill connections for those interested in skills-based hiring.

Within the O*NET ecosystem, these indices would have several potential applications, such as: (a) individual knowledges and essential/transferrable skills for an occupation in O*NET Online could be tagged with an indicator or rating of the potential for AI to exhibit the required level of knowledge/skill for the occupation in question; (b) the knowledge and essential/transferrable skill level ratings could be aggregated to the occupation-level—using either unweighted or importance-weighted averages—to yield an alternative, knowledge and skill-centric occupation-level indexes; (c) the software and vacancy-based skill indices could be aggregated to the occupation-level to indicate how much of a nexus there is between AI and software and vacancy-sourced skills for an occupation, and lastly (d) the indexes could contribute to an ensemble AI impact composite that synthesizes evidence across multiple indices.

Task Performance-Focused Indices (Indices 5–8)

The next four indices shift focus from the determinants of performance to one of its major components: task performance. Task performance encompasses the core technical activities and responsibilities that define what it means to do a job well, as reflected in the behaviors most directly documented in formal job descriptions and O*NET task statements. These indices are organized along two sub-dimensions: unit of analysis (job-element vs. occupation-level) and impact type (automation vs. augmentation), yielding a two-by-two structure that enables meaningful comparisons both within and across units of analysis and impact type.

Index 5 is a job task-level index focused on evaluating whether the current state of AI technology can automate the performance of a given job task. This index would bear similarity to several AI impact approaches in the literature, many of which assign AI automation potential scores to individual tasks. *Index 6* is the augmentation-focused counterpart to Index 5. Index 6 evaluates whether current AI technology can augment human performance of each task; that is, whether AI can serve as a co-pilot that enhances human capability on the task rather than replacing the human altogether. The parallel construction of Indices 5 and 6 is deliberate: presenting both side-by-side for the same task allows users to appreciate that automation and augmentation are not simply opposite ends of a single spectrum but distinct forms of AI impact with different implications for workforce planning and worker experience. Within O*NET Online, individual tasks can be tagged with automation and augmentation indicators, offering users a granular view of which specific task demands within an occupation are most susceptible to AI-driven automation and augmentation.

Indices 7 and 8 move the focus up a unit of analysis to offer occupation-level indices of automation and augmentation potential (respectively), derived not from aggregating task-level scores but from directly prompting LLMs to evaluate the occupation as a whole. This is an important design choice because job performance may not simply be the sum of its elemental parts. Specifically, aggregated individual task scores at the occupation-level may miss important interdependencies among tasks, sequencing requirements, and contextual factors that bear on whether an occupation can be automated as a whole. Together, Indices 7 and 8 provide a holistic, task-performance-centric view of AI's potential impact at the occupation-level that could help populate a new AI impact area of the O*NET Content Model, and both can be compared against their task-level aggregate counterparts (Indices 5 and 6) and against established occupation-level AI impact indices from prior studies.

Contextual Performance-Focused Indices (Indices 9–12)

One of the most distinctive features of the proposed suite, and one that most sharply differentiates it from existing approaches in the literature, is its explicit attention to contextual performance. While task performance captures the core technical activities of a job, contextual performance encompasses the behaviors that support the broader organizational, social, and psychological environment in which work occurs, such as interpersonal citizenship behaviors, organizational commitment, helping and cooperating with others and demonstrating effort that goes above and beyond ones' formally proscribed role (Borman & Motowidlo, 1993, 1997; Campbell & Wiernik, 2015). These behaviors are critical to organizational functioning and are a long-established component of job performance recognized by organizational researchers. Despite their centrality to modern definitions of job performance, they are largely, if not entirely, absent from existing AI impact frameworks. By offering a dedicated set of contextual performance-focused indices, we acknowledge that any assessment of AI's impact on work that is restricted to task performance is, by definition, incomplete from the perspective of the scientific literature on job performance.

A practical challenge with offering contextual, performance-focused AI impact indices is that O*NET does not currently maintain lists of occupation-relevant contextual performance behaviors (e.g., volunteering to mentor or onboard junior staff) for each occupation (in contrast to its maintenance of task statements). Accordingly, to form these indices, we would first prompt one or more LLMs to generate occupationally relevant contextual performance behaviors for each occupation, then use those behaviors as inputs to prompts designed to produce the AI impact index of interest.

Index 9 is a contextual behavior-level index evaluating whether current AI can automate the performance of occupation-relevant contextual performance behaviors. *Index 10* is the augmentation-focused counterpart to Index 9. Together, Indices 9 and 10 would allow the O*NET Program to present a profile of automation and augmentation potential separately for contextual behaviors, enabling comparisons to the analogous task-performance findings and illuminating potentially important differences in AI impact conclusions for task vs. contextual performance.

Index 11 is an occupation-level index evaluating whether current AI can automate the performance of an occupation as characterized by its description and collective set of contextual performance behaviors. *Index 12* is the augmentation-focused counterpart to Index 11, using the same inputs and process with an augmentation-oriented rating scale. Indices 11 and 12 can potentially be used to populate the new AI impact area of the O*NET Content Model alongside the task-performance occupation-level indices, and comparisons among these indices can shed light on the degree to which AI's impact on an occupation differs depending on whether one focuses on task or contextual performance within an occupation.

Overall Job Performance-Focused Indices (Indices 13–16)

The final four indices offer the broadest vantage point on AI impact among the suite's indices, focusing on overall job performance rather than its individual components. These indices are designed to provide integrative, occupation- and vacancy-level assessments that reflect AI's potential impact on the full scope of what it means to perform a job, encompassing all demands simultaneously. *Index 13* is an occupation-level automation index that evaluates whether current AI can automate an occupation's overall performance based on its description and the combined set of job tasks and contextual performance behaviors. By incorporating inputs from both performance components as inputs, Index 13 offers a more comprehensive occupation-level automation picture than can be obtained from Indices 7 or 11 alone. *Index 14* is the augmentation-focused counterpart to Index 13, using identical inputs and process but with an augmentation-oriented scale. Taken together, Indices 13 and 14 represent the suite's most comprehensive occupation-level indicators, capturing AI's potential to automate or augment the full set of performance tasks/behaviors that define work in an occupation rather than any single component.

Index 15 introduces a distinct unit of analysis (the job vacancy) and offers an automation-focused index derived from actual job postings rather than standardized O*NET occupation profiles. Index 15's primary value is complementing the O*NET-occupation data-based indices, as they are based on more granular vacancy data rather than standardized occupation-level data. As such, this index would afford a window into within-occupation heterogeneity in automation potential across different positions and employers, which is a variation that occupation-level indices, by design, cannot capture. *Index 16* is the augmentation-focused counterpart to Index 15. Together, Indices 15 and 16 round out the suite by providing a vacancy-grounded perspective on AI impact that can be used both as a substantive lens on within-occupation variation and as an evaluative comparison point against native occupation-level indices based on standardized occupational data.

Updating AI Impact Index Ratings Over Time

A key advantage of the AI impact described above is that all of them could be readily refreshed over time. For Indices 1-3 and 5-8, the primary driver of refresh needs will be updates to O*NET's knowledge, skill, and task content for specific occupations or the introduction of new occupations. Because our envisioned generation process is prompt-based and relies on O*NET's own data as inputs, refresh can be executed by re-running the relevant prompts whenever O*NET content is updated, without requiring new human rating studies or extensive manual intervention. For Indices 9–14, refresh will additionally require re-running the contextual behavior generation process when occupation content changes significantly. For Indices 4, 15, and 16, refresh is driven by the availability of new vacancy data, and the process can, in principle, be run on a rolling basis as new postings become available, offering a more continuously updated vacancy-level perspective on AI impact.

Of course, all of the indices above can and should be refreshed regularly to reflect the fact that AI technology and its capabilities in the world of work are constantly evolving. As such, even without changes to job elements and occupation inputs into the indices, the indices should be regularly refreshed to account for advances in AI technology and improvements to the LLMs used to generate them.

Regardless, should the O*NET Program decide to adopt one or more of the indices proposed above, it will be important to establish a policy for refreshing and re-evaluating the indices over time. Presumably, replacing any of the LLMs used to generate the initial evaluation of the ratings (e.g., due to deprecation of the original models, or interest in evaluating more powerful models), will have to be coupled with an evaluation of those new models to ensure the meet or exceed the quality of the models initially used and evaluated for producing these indices (e.g., see Putka et al., 2025 for an example of such a LLM-re-evaluation plan in the context of maintaining O*NET Work Style ratings).

One additional recommendation we have for the O*NET Program in this regard is to maintain AI impact index data in a way that allows users of the O*NET ecosystem to understand how AI's impact on work evolves over time (e.g., by displaying visualizations of changes in AI impact indices for an occupation across yearly quarters). Given how fast AI capabilities are evolving, with regular refreshes of these indices, the O*NET Program could position itself well as a central source for understanding how AI's impact on the world of work is shifting over time, wrapped around a well-established gold-standard occupational information framework.

Proposed AI Impact Index Evaluation Strategy

A desired characteristic of the proposed suite of AI impact indices is that they can be evaluated for quality. That is, there should be principled means of assessing whether each index does what it is intended to do and builds an evidentiary basis for it. Though not detailed in this report, at a high level, we would recommend an evaluation strategy that focuses at least on two broad areas: (1) evaluating the reliability of the AI impact index ratings; and (2) establishing a nomological net for the proposed suite of indices and using that as a basis for evaluating expected patterns of convergence and discrimination between (a) proposed indices and external data (i.e., AI impact from indices from past work, independent expert ratings), and/or (b) the proposed set of indices themselves.

Reliability Evaluation

As described earlier, we envision the final rating for any given AI impact index to be based on a composite of ratings from multiple LLMs. For each LLM, we would propose running the given prompt multiple times and averaging the scores before assembling ratings across LLMs. Basing the AI impact ratings on multiple models and runs reduces the idiosyncratic perspective of any single LLM or run stemming from the stochastic nature of the generative process. From a reliability-estimation perspective, this would allow us to estimate and calibrate "model-specific" and "run-specific" sources of error in ratings, and in turn estimate a multifaceted estimate of AI impact reliability. Specifically, it would allow for the indexing of (a) error stemming from inconsistency in ratings across LLMs (i.e., model-specific error), (b) error stemming from inconsistency in ratings across occasions on which a given model was run (i.e., run-specific error).

Establishing a Nomological Net for the Proposed Suite of Indices

To aid in the evaluation of the proposed indices, we would also recommend developing an AI impact-centric nomological net, that is, a theoretically grounded set of expectations about how the proposed indices should relate to one another and to established indices from prior research, and in some cases, independent expert ratings. Specifying such a net in advance serves two important functions. First, it provides a principled basis for evaluating the proposed indices rather than relying on post-hoc rationalization of whatever patterns emerge. Second, it situates the proposed suite within the broader cumulative literature on AI impact, enabling the O*NET Program to position its indices in relation to, rather than in isolation from, the considerable body of prior work in this space.

For proposed AI impact indices with conceptual analogs in the prior AI impact literature, a convergent validity evaluation would involve comparing index values to corresponding estimates from established prior studies to assess the extent of correlational similarity. By "corresponding estimates," we mean AI impact indices from prior work that share conceptual similarities with the AI impact index we propose here (e.g., the pair of indices both target O*NET tasks and evaluate potential for AI automation).

In addition to examining convergence between the proposed AI impact indices and external data, validity evidence can also be obtained by examining patterns of convergence and discrimination among the proposed indices themselves. Specifically, one could analyze correlations among indices through a Multitrait-Multimethod (MTMM) correlation lens to evaluate patterns of convergence and discrimination among the proposed set of AI impact indices. The purpose of doing so is to evaluate evidence of convergent and discriminant validity for the different AI impact indices (Campbell & Fiske, 1959). In this context, convergent validity would be indicated by high correlation among indexes targeting the same outcome (e.g., task performance, contextual performance) expressed at or aggregated to the same unit of analysis (e.g., occupation-level) – such correlations could be framed as Monotrait-Heteromethod (MT-HM) correlations in the classical MTMM sense. Discriminant validity and freedom from "common method" variance are indicated by relatively lower correlations among indices that target different outcomes (e.g., task vs. contextual performance) but share an approach in common (e.g., focus on automation at the occupation-level, focus on augmentation at the task/behavior level) – such correlations could be framed as HT-MM correlations in the classical MTMM sense. The lowest correlations should be among those indices that share neither a target outcome nor approach (i.e., different automation/augmentation focus, different focal unit of analysis). Among the suite of proposed AI impact indices, we would expect to see several differ type of patterns in correlations should the indices be functioning as intended.

Potential Future Considerations

In the sections above, we outlined a suite of AI impacts for the O*NET Program to consider, grounded in past research and modern perspectives on job analysis and job performance. Nevertheless, other possibilities could be considered in the future, depending on stakeholder interests. For example, in proposing the set of indices above, we deliberately expanded beyond task performance to include contextual performance. However, there are other domains of performance that may be of interest to the O*NET Program, most notably AI's impact on *adaptive performance* (Baard et al., 2013; Dorsey et al., 2017; Pulakos et al., 2020). Adaptive performance can be defined as a "change in response to an altered situation or the behavioral outcome of the adaptation process" (Dorsey et al., 2017, p. 440). The reason we decided not to formally include AI impact indices that target adaptive performance in our proposed suite is that there are varying perspectives in the research literature on whether concepts of adaptation at work is best framed as (a) a set of individual difference constructs (e.g., ability, capacity, and willingness to adapt), (b) a process, or (c) a job performance domain (as defined above) (Dorsey et al., 2017). That being said, to the extent the O*NET Program (or its stakeholders) has an interest in defining adaptation from a job performance perspective, it may want to consider developing an additional pair of indices (i.e., automation and augmentation) targeting adaptive performance comparable to what we have proposed above for contextual performance. If the decision is to pursue this avenue, the prevailing taxonomies of adaptive performance should be considered to aid in the process of identifying/generating examples of adaptive performance behavior for each occupation (e.g., Pulakos et al, 2020).

Another potential future consideration is revisiting how the "motivation" performance determinant may manifest in determining AI's impact on job performance. Earlier, we noted that we excluded targeting motivation in the proposed set of AI impact indices because it has largely been defined in terms of *humans'* choice to direct, allocate, and persist with effort (Campbell et al., 1993). However, there may be analogs of these concepts for AI agents (Shao et al., 2025). For example, *direction* could be linked to AI agent objective functions and policy optimization, *allocation* could be linked to how many compute resources are allocated to an AI agent (as well as attention weights underlying the model that drives the agent's functioning), and *persistence* could be linked to convergence criteria and stopping rules for AI agents iterating on a task. Then again, many of the links above could be framed as choices not on the part of the AI, but rather choices on the part of the human developer of the AI that have implications for the agent's direction, allocation, and persistence with effort – so would it truly be motivation on the part of the AI, or simply be the AI behaving as instructed/designed? All this being said, fleshing out an AI agent-centric model of job performance that includes AI agent DK, PKS, and motivation performance determinants was beyond the scope of the current effort, but we raise it as a future possibility as AI agents begin to mimic volitional human behavior.

Lastly, we recognize that we have proposed developing many potential AI indices. While offering a justification and use for each index in O*NET, as well as clarifying how they tie back to a contemporary model of job performance, we stopped short of how an O*NET user may use these indices *as a set* to inform decision making. For example, if an O*NET user or policy maker aims to use the set of indices to inform a decision about an occupation or broader policy, how might the indices be combined into a clear decision framework? Should the O*NET Program move forward developing these indices, more thought would need to be given to how they might be used in combination to inform various types of decisions or policy inferences a user may wish to draw. We stopped short of doing that here as we envision that the final set of indices that survive any development and evaluation process may look slightly different (potentially reduced) relative to the set above, and any use of them in combination would in part depend on specific use cases and how they empirically relate to one another – all of which is yet to be determined.

Closing Thoughts

This report provides a structured review of methods adopted to assess AI's impact on work, along with recommendations for developing and maintaining a suite of job-analysis-centric AI impact indices for the O*NET ecosystem. Beginning with a set of core design principles that embrace a job analytic perspective and modern theories of job performance, we have proposed a suite of 16 distinct indices, grounded in a contemporary model of job performance and its determinants.

The proposed suite offers several distinctive strengths relative to the existing AI impact work. Its explicit grounding in established job performance theory—encompassing task performance, contextual performance, and their proximal determinants (declarative knowledge, procedural knowledge, and skill) represents a meaningful departure from the predominantly task-centric approaches that have characterized prior work. Like past work, it draws a systematic distinction between automation and augmentation, acknowledging that these are fundamentally different phenomena with distinct implications for workers, organizations, and policy. Its multi-level architecture allows AI impact data to be presented and analyzed at the level of individual tasks, knowledge, skills, vacancies, and occupations, supporting a wider range of potential use cases than any single-level approach can accommodate. Lastly, its explicit commitment to evaluability, as evidenced by recommendations for specifying a nomological net and using that as a foundation for subsequent validation, provides a principled foundation for evaluating index quality. As highlighted in our overview of proposed AI impact indices, the set of indices could be used in several different ways in the O*NET ecosystem, ranging from adding AI-impact-related details to an occupation's existing profile online to creating and maintaining an AI-impact-centric part of the O*NET Content Model and database to facilitate future users' decision making, applications, and research.

Developing, evaluating, and maintaining the proposed suite will require a meaningful investment and development of partnerships on the O*NET System's part. That being said, with sustained commitment and careful execution, the proposed suite has the potential to establish O*NET and DOL as a key provider of gold standards for AI impact data in the United States—delivering rigorous, transparent, and regularly updated information about the relationship between AI and work that can help serve the needs of workers, employers, researchers, and policymakers for years to come.

References

References marked with a single asterisk (*) represent core studies that we reviewed that employed O*NET data in some manner for purposes of assessing AI impact. References marked with a double asterisk (**) represent additional publications that, while not included in our set of core studies that we reviewed and mapped to Figure 1, nonetheless also employed O*NET data in some capacity.

*Agarwal, P., Agarwal, H., & Rana, A. (2025). *What work is AI actually doing? Uncovering the drivers of generative AI adoption.* arXiv. https://doi.org/10.48550/arXiv.2510.23669

*Ahmadi, M., Kheslat, N. K., & Akintomide, A. (2024). *Generative AI impact on labor market: Analyzing ChatGPT's demand in job advertisements*. arXiv. https://doi.org/10.48550/arXiv.2412.07042

Alekseeva, L., Azar, J., Giné, M., Samila, S., & Taska, B. (2021). The demand for AI skills in the labor market. *Labour Economics*, *71*, 102002. https://doi.org/10.1016/j.labeco.2021.102002

*Appel, R., Massenkoff, M., McCrory, P., McCain, M., Heller, R., Neylon, T., & Tamkin, A. (2026). *Anthropic Economic Index report: Economic primitives*. Anthropic Economic Research. https://www.anthropic.com/research/anthropic-economic-index-january-2026-report

Baard, S. K., Rench, T. A., & Kozlowski, S. W. J. (2013). Performance adaptation: A theoretical integration and review. *Journal of Management, 40*(1), 48-99. https://doi.org/10.1177/0149206313488210

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71–98). Jossey-Bass.

Borman, W. C., & Motowidlo, S. J. (1997). Task performance and contextual performance: the meaning for personnel selection research. *Human Performance, 10*(2) 99-109. https://doi.org/10.1207/s15327043hup1002_3

Brannick, M. T., Levine, E. L., & Morgeson, F. P. (2007). *Job and work analysis: Methods, research, and applications for human resource management* (2nd ed.). Sage Publications, Inc.

*Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings, 108,* 43-47. https://doi.org/10.1257/pandp.20181019

**Cai, A., YeckehZaare, I., Sun, S., Charisi, V., Wang, X., Imran, A., ... & Malone, T. W. (2026). *Where can AI be used? Insights from a deep ontology of work activities.* arXiv. https://doi.org/10.48550/arXiv.2603.20619

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81-105. https://doi.org/10.1037/h0046016

Campbell, J., McCloy, R., Oppler, S., & Sager, C. (1993). A theory of performance. In N. Schmitt, & W. Borman (Eds.), *Personnel Selection in Organizations* (pp. 35-70). Jossey-Bass.

Campbell, J., & Wiernik, B. M. (2015). The modeling and assessment of job performance. *Annual Review of Organizational Psychology and Organizational Behavior, 2*, 47-74.

Carnevale, A., Jayasundera, T., & Repnikov. D. (2014). *Understanding online job ads data*. Georgetown University Center on Education and the Workforce. https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM.Tech_.Web_.pdf

**Chen, W. X., Srinivasan, S., & Zakerinia, S. (2025). *Displacement or complementarity?: The labor market impact of generative AI.* (Working Paper). https://www.hbs.edu/ris/download.aspx?name=25-039.pdf

*Chopra, A., Bhattacharya, S., Salvador, D., Paul, A., Wright, T., Garg, A., ... & Balaprakash, P. (2025). *The Iceberg Index: Measuring skills-centered exposure in the AI economy*. arXiv. https://doi.org/10.48550/arXiv.2510.25137

Dorsey, D. W., Cortina, J. M., Allen, M. T., Waters, S. D., Green, J. P., & Luchman, J. (2017). Adaptive and citizenship-related behaviors at work. In Farr, J. L, & Tippins, N. T. (Eds.), *Handbook of employee selection* (2nd ed., pp. 448-475). Routledge.

*Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, *384*(6702), 1306-1308. https://doi.org/10.1126/science.adj0998

**Felten, E. W., Raj, M., & Seamans, R. (2018). A method to link advances in artificial intelligence to occupational abilities. *AEA Papers and Proceedings, 108*, 54–57. https://doi.org/10.1257/pandp.20181021

*Felten, E. W., Raj, M., & 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. https://doi.org/10.1002/smj.3286

*Fossen, F. M., & Sorgner, A. (2019). *New digital technologies and heterogeneous employment and wage dynamics in the United States: Evidence from individual-level data* (No. 12242). IZA Discussion Papers. https://docs.iza.org/dp12242.pdf

Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., Moro, E., Wang, D., Youn, H., & Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences, 116*(14), 6531–6539. https://doi.org/10.1073/pnas.1900949116

**Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation?. *Technological Forecasting and Social Change*, *114*, 254-280. https://doi.org/10.1016/j.techfore.2016.08.019

Gael, S. (1988). *The job analysis handbook for business, industry, and government.* John Wiley and Sons.

**Garcia-Lluis Valencia, A. (2026). *ITEA framework: A multidimensional system for measuring occupational exposure to algorithmic expropriation under the agentic AI regime.* https://github.com/Aval22/ITEA

*Hampole, M., Papanikolaou, D., Schmidt, L. D., & Seegmiller, B. (2025). *Artificial intelligence and the labor market* (No. w33509). National Bureau of Economic Research. https://doi.org/10.3386/w33509

*Handa, K., Tamkin, A., McCain, M., Huang, S., Durmus, E., Heck, S., ... & Ganguli, D. (2025). *Which economic tasks are performed with AI? evidence from millions of Claude conversations*. arXiv. https://doi.org/10.48550/arXiv.2503.04761

Hansen, M. C., Norton, J. J., Gregory, C. M., Meade, A. W., & Thompson, L. F. (2014). *O*NET work activities project technical report*. National Center for O*NET Development. https://www.onetcenter.org/reports/DWA_2014.html

Hui, X., Reshef, O., & Zhou, L. (2024). The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. *Organization Science*, *35*(6), 1977-1989. https://doi.org/10.1287/orsc.2023.18441

Hurtz, G. M., & Wright, C. W. (2012). Designing work descriptions to maximize the utility of employee recruitment efforts. In M. A. Wilson, W. Bennet, S. Gibson, & G. M. Alliger (Eds.), *The handbook of work Analysis: Methods, systems, applications and science of work measurement in organizations* (pp. 347-379). Routledge.

International Economic Development Council. (2025). *Artificial intelligence impact on labor markets: Literature review*. Economic Development Research Partners. https://www.iedconline.org/edrp-reports/ai/artificial-intelligence-impact-on-labor-markets/

International Organization for Standardization (2022). *Information technology – Artificial intelligence – Artificial intelligence concepts and terminology* (ISO/IEC 22989:2022). https://www.iso.org/standard/74296.html

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*(4), 657–690. https://doi.org/10.1037/0021-9010.74.4.657

*Kochhar, R. (2023). *Which U.S. workers are more exposed to AI on their jobs?* Pew Research Center. https://www.pewresearch.org/social-trends/2023/07/26/which-u-s-workers-are-more-exposed-to-ai-on-their-jobs/

Lancaster, V., Mahoney-Nair, D., & Ratcliff, N. (2019). *Technology report: Review of Burning Glass job-ad data*. University of Virginia Biocomplexity Institute. https://libraopen.library.virginia.edu/server/api/core/bitstreams/1e9486ad-98a2-4d61-9c14-2602a38fc0b9/content

Mäkelä, E., & Stephany, F. (2025). *Complement or substitute? How AI increases the demand for human skills*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5153230

**Massenkoff, M, & McCrory, P. (2026). *Labor market impacts of AI: A new measure and early evidence.* Anthropic Economic Research. https://www.anthropic.com/research/labor-market-impacts

McCloy, R. A., Campbell, J. P., & Cudeck, R. (1994). A confirmatory test of a model of performance determinants. *Journal of Applied Psychology, 79*(4), 493–505. https://doi.org/10.1037/0021-9010.79.4.493

McCormick, E. J. (1979). *Job analysis: Methods and applications.* American Management Associations.

Motowidlo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance, 10*(2), 71–83. https://doi.org/10.1207/s15327043hup1002_1

National Center for O*NET Development (2026a). *O*NET 30.2 database*. Author. Retrieved April 13, 2026, from https://www.onetcenter.org/database.html

National Center for O*NET Development. (2026b) *Office of Management and Budget (OMB) clearance*. O*NET Resource Center. Retrieved April 14, 2026, from https://www.onetcenter.org/ombclearance.html

OECD (2019). *Recommendation of the council on artificial intelligence*. OECD Publishing, Paris. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

*OECD (2025). *OECD AI capability indicators technical report.* OECD Publishing, Paris. https://doi.org/10.1787/9cdb3dd1-en

**Pizzinelli, C., Panton, A., Tavares, M. M., Cazzaniga, M., & Li, L. (2023). *Labor market exposure to AI: Cross-country differences and distributional implications* (IMF Working Paper No. WP/23/216). International Monetary Fund. https://doi.org/10.5089/9798400254802.001

*PricewaterhouseCoopers (PwC). (2025). *The fearless future: Global AI jobs barometer 2025* (Methodology Appendix). https://www.pwc.com/gx/en/issues/artificial-intelligence/job-barometer/2025/report.pdf

Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2020). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85*(4), 612-624. https://psycnet.apa.org/doi/10.1037/0021-9010.85.4.612

Putka, D. J., Liu, J., Wu, F., Burke, M., & Lewis, P. (2025). *Using a hybrid artificial intelligence-expert methods to develop Work Style ratings for the O*NET database*. National Center for O*NET Development. https://www.onetcenter.org/reports/Hybrid_AI_Ratings.html

**Ryseff, J. (2026). *AI's potential impact on the Army Officer Corps.* Special Competitive Studies Project. https://www.scsp.ai/ai-potential-impact-on-the-army-officer-corps/

Sajadieh, S., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Santarlasci, L., Pava, J., Maslej, N., Altman, R., Brynjolfsson, E., Brodley, C., Clark, J., Dignum, V., Kumar, V., Landay, J., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Tabassi, E., Wald, R., Walsh, T., & Weld, D. (2026). *The AI Index 2026 annual report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. https://hai.stanford.edu/assets/files/ai_index_report_2026.pdf

Samoili, S., López Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F., and Delipetrev, B. (2020). *AI Watch. Defining artificial intelligence. Towards an operational definition and taxonomy of artificial intelligence* (EUR 30117 EN). Publications Office of the European Union, Luxembourg. https://doi.org/10.2760/382730

Schmidt, F. L. (2014). A general theoretical integrative model of individual differences in interests, abilities, personality traits, and academic and occupational achievement: A commentary on four recent articles. *Perspectives on Psychological Science*, *9*(2), 211–218. https://www.jstor.org/stable/44289978

*Shao, Y., Zope, H., Jiang, Y., Pei, J., Nguyen, D., Brynjolfsson, E., & Yang, D. (2025). *Future of work with AI agents: Auditing automation and augmentation potential across the US workforce*. arXiv. https://doi.org/10.48550/arXiv.2506.06576

Sigelman, M., Francis, B., Rao, S., & Guilford, G. (2026). *Beyond the binary: How automation and augmentation are combining to reshape work* (Working Paper). Burning Glass Institute. https://www.burningglassinstitute.org/research/beyondthebinary

Theofanos, M., Choong, Y., & Jensen, T. (2024). *AI use taxonomy: A human-centered approach.* National Institute of Standards and Technology, U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.200-1

*Tomlinson, K., Jaffe, S., Wang, W., Counts, S., & Suri, S. (2025). *Working with AI: Measuring the applicability of generative AI to occupations*. arXiv. https://doi.org/10.48550/arXiv.2507.07935

U.S. Congress. (2021). *National Artificial Intelligence Initiative Act of 2020* (Pub. L. No. 116–283, div. E, 134 Stat. 4523). https://www.congress.gov/bill/116th-congress/house-bill/6216

U.S. Department of Labor (2026). *The U.S. Department of Labor's artificial intelligence literacy framework* (Training and Employment Notice No. 07-25). Employment and Training Administration. https://www.dol.gov/agencies/eta/advisories/ten-07-25

*Webb, M. (2019). *The impact of artificial intelligence on the labor market* (Working Paper). https://doi.org/10.2139/ssrn.3482150

Wilson, M. A., Bennet, W., Gibson, S., & Alliger, G. M. (Eds.) (2012). *The handbook of work analysis: Methods, systems, applications and science of work measurement in organizations*. Routledge.

Wilson, M. A. (2007). A history of job analysis. In. Koppes, L. L. (Ed.), *Historical perspectives in industrial and organizational psychology* (pp. 219-241). Lawrence Erlbaum Associates.

Workday, & Access Partnership. (2026). *Labor data readiness in the age of AI: Building the foundations for smarter workforce planning*. https://accesspartnership.com/wp-content/uploads/2026/01/Workday-Labor-Data.pdf

Appendix A: Overview of AI Definitions, Frameworks, and Taxonomies

Due to the rapidly increasing impact of AI on the labor market across multiple sectors (see Frank et al., 2019; Hui et al., 2024; International Economic Development Council, 2025; Massenkoff & McCrory, 2026; Ryseff, 2026), as well as growing efforts to respond to the increasing prominence of AI within the workforce (Sajadieh et al., 2026; U.S. Department of Labor, 2026; Workday, & Access Partnership, 2026), we also scanned the literature for prominent AI definitions and taxonomies. Doing so helped ensure we had an understanding of how AI has historically been framed and conceptualized from a taxonomic perspective at the outset of this effort. Furthermore, we viewed this information as potentially useful to the O*NET Program should it move to implement the recommendations outlined in this report (e.g., as an aid in developing an operationalization of AI for use in AI impact-related prompts).

The information reviewed included general-purpose definitions, such as those articulated in the National Artificial Intelligence Initiative Act (U.S. Congress, 2021) and OECD (2019, 2025). We also examined other prominent frameworks, such as the (1) National Institute of Standards and Technology's (NIST) AI Use Taxonomy (Theofanos et al., 2024), which classifies how AI contributes to tasks across 16 activities (e.g., content creation, decision making, prediction) (2) the European Commission's AI Taxonomy (Samoili et al., 2020), which taxonomizes AI activities across key domains (e.g., reasoning, planning, communication) and subdomains (e.g., knowledge representation, searching, natural language processing), and (3) the International Organization for Standardization's (ISO; 2022) list of AI concepts (e.g., neural networks, AI trustworthiness, AI system lifecycles) and terminologies (e.g., generative AI, bias).
For example, the OECD AI Capability Indicators (OECD, 2025) constitute a human-centered, capability-based taxonomy that assesses AI's ability to replicate or approximate human cognitive and physical capacities. Specifically, the OECD framework organizes AI into nine capability domains (e.g., language, problem-solving, creativity, social interaction) and places each on a five-level scale representing current AI capabilities. Accordingly, this approach is designed for policy and planning with the goal of helping stakeholders understand what AI can do relative to humans and how those capabilities may evolve over time. The NIST AI Use Taxonomy (Theofanos et al., 2024) has a slightly different focus, conceptualizing AI in terms of its role in human tasks. Instead of categorizing AI by capability or domain, it defines 16 "AI use activities" that describe how AI contributes to outcomes in human-AI interactions, independent of specific techniques. Some examples of these include content creation ("The AI system assists by… generating new artifacts such as video, narrative, software code, synthetic data" page 4) and process automation ("The AI system assists by… performing repetitive tasks, removing bottlenecks, reducing errors and loss of data, and increasing efficiency of a process" page 4). Additionally, the European Commission's AI taxonomy (Samoili et al. 2020) organizes AI applications into core technical domains (e.g., reasoning, learning, communication, and perception) and several subdomains, including knowledge representation, machine learning, computer vision, robotics, and automation. This taxonomy also includes a large keyword set for each domain/subdomain, which researchers and policymakers can use to map and monitor AI activity across the workforce at a more granular level. Rather than focusing specifically on what AI can do, this framework emphasizes the broader AI ecosystem, making it especially useful for classification, benchmarking, and tracking technological development, and providing a common foundation for what constitutes AI.

Similar to the European Commission AI taxonomy, which contains an underlying set of key AI terms, we reviewed other sources that contained lists/definitions of specific AI applications, as this can also be informative for understanding how AI is operationalized. For example, the

International Organization for Standardization (ISO) (2022) provides a list of over 100 keywords (e.g., AI agent, narrow AI, semantic computing, transfer learning, neural networks) that are located within several larger AI categories (e.g., Types of AI, machine learning algorithms, data concepts, AI trustworthiness concepts). Additionally, many studies have used AI keyword lists to inform their estimates of AI impact. For example, when identifying jobs that require AI skills, PwC (2025) identified nearly 400 such skills (e.g., cognitive computing, robotic systems, zero-shot learning, fuzzy logic). Moreover, in developing their task-level estimate of AI impact, Hampole et al. (2025) examined over 30 AI terms within job descriptions (e.g., deep learning, LLM, computer vision, generative pre-trained transformer) while Alekseeva et al. (2021) used a list of over 70 skills via Burning Glass Technologies (e.g., AI chatbot, object recognition, support vector machines, TensorFlow) to help identify AI-relevant vacancies. Moreover, Mäkelä and Stephany (2025) also used a large list of over 100 AI skills (e.g., gradient boosting, feature extraction, predictive modeling, pattern recognition) to identify AI-related job vacancies.

Collectively, these taxonomies and studies employing AI keyword lists highlight the broad ways in which AI can be defined and conceptualized. This is important to keep in mind when assessing AI impact, as it indicates that the conclusions drawn from any study of AI impact will be partially a function of how AI is defined and operationalized. Below, we provide a summary table of these frameworks for those interested in learning more.

Table A.1. Summary of Select AI Taxonomies and Frameworks

Source	Resource	Summary
NIST	AI Use Taxonomy	This taxonomy adopts a human-centered approach to classify how AI systems contribute to specific outcomes. It employs 16 distinct AI use activities (e.g., content creation, decision-making, and detection) that describe how an AI system either augments or replaces human effort for a given task. By organizing complex human-AI interactions into these activities, this taxonomy provides a framework for evaluating trustworthy AI. A key aim of this taxonomy is to facilitate an improved understanding of human-AI tasks to ensure optimal outcomes and better AI risk management. (Theofanos et al., 2024) See: https://doi.org/10.6028/NIST.AI.200-1
OECD	AI Capability Indicators	These indicators provide an evidence-based framework for policy makers to assess AI progress by comparing it directly to human abilities. The framework contains nine dimensions of human capabilities (i.e., language; social interaction; problem solving; creativity; metacognition and critical thinking; knowledge, learning and memory; vision; manipulation; and robotic intelligence). Each indicator is measured on a five-point scale, ranging from basic solved challenges (1) to full human equivalence (5). This system will ideally help stakeholders anticipate AI's impact by identifying gaps between current and hypothetical future AI capabilities. (OECD, 2025) See: https://doi.org/10.1787/9cdb3dd1-en
EC	AI Taxonomy	This taxonomy provides an operational definition of AI to monitor its development and impact. It employs a multidimensional structure and is organized into core domains (e.g., reasoning, planning, learning, communication, and perception) and various subdomains (e.g., knowledge representation, robotics and automation, multi-agent systems). The taxonomy also denotes relevant keywords for each domain (e.g., feature extraction, network intelligence, self-driving car) The goal is for this taxonomy to be iteratively updated to enable tracking rapid developments in AI over time. (Samoili et al. 2020) See: https://doi.org/10.2760/382730
ISO	AI Concepts and Terminology	This standard establishes a set of concepts and terminology for AI to help facilitate clear communication among different AI stakeholders. It defines AI as a discipline and an AI system as an engineered system that generates outputs like content, predictions, or decisions based on human-defined objectives. It provides a detailed overview of the AI system life cycle, ranging from inception and development to verification and retirement. Additionally, it categorizes critical components of AI systems (e.g., data processing, machine learning) while defining AI key trustworthiness properties such as transparency, and explainability. (International Organization for Standardization, 2022) See: https://www.iso.org/standard/74296.html

Note. NIST = National Institute of Standards and Technology. OECD = Organisation for Economic Co-operation and Development. EC = European Commission. ISO = International Organization for Standardization.