Asia-Pacific
Economic Cooperation

Advancing Free Trade
for Asia-Pacific **Prosperity**

Big Data for the Labor Market: Sources, Uses and Opportunities

ISSUES PAPER No. 13

APEC Policy Support Unit

December 2021

Prepared by:
Julia Nitschke, Layla O’Kane, Bledi Taska, and Nyerere Hodge
Emsi Burning Glass
66 Long Wharf, 2nd Floor, Boston, Massachusetts
United States of America 02110
Website: www.burning-glass.com

Emmanuel A. San Andres and Andre Wirjo
Asia-Pacific Economic Cooperation Policy Support Unit
Asia-Pacific Economic Cooperation Secretariat
35 Heng Mui Keng Terrace
Singapore 119616
Tel: (65) 6891-9600 Fax: (65) 6891-9690
Email: psugroup@apec.org Website: www.apec.org

Produced for:
Asia-Pacific Economic Cooperation

APEC#221-SE-01.19

The views expressed in this paper are those of the authors and do not necessarily represent those of APEC Member Economies. Terms such as ‘national’, nation’, ‘country’ and ‘state’ used in the text are for purposes of this paper and do not imply the political status of any APEC Member Economy. This project benefited from a research grant provided by The Asia Foundation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF BOXES

EXECUTIVE SUMMARY

The COVID-19 pandemic has drastically changed the labor market and exacerbated trends that were already on the rise. Rapid technological change characterized by digitalization and automation of jobs has become even more pronounced. Amid this economic crisis, access to real-time labor market data has become even more pivotal in shaping policy.

Working under the APEC Policy Support Unit (PSU) project on 'Filling APEC's Data Gaps to Address Future of Work Challenges', Emsi Burning Glass (EBG) has conducted a mixed methods study of existing labor market data sources, and integration of new data into labor market systems, and assessed opportunities for public–private partnerships.

Research questions

- What gaps exist in current labor market information in APEC economies, particularly with regard to digital jobs and skills?
- What new data sources are available, such as big data analysis of job postings and career histories on digital labor market platforms? How can these new data sources be integrated into current labor market systems?
- What opportunities for public–private partnerships exist and how have they succeeded in the past?

Key findings

- Big data can help augment traditional sources of data in providing real-time analyses and are especially useful in times of economic shock.
- Big data allow for analyses such as calculating skill premia, understanding skill adjacencies and creating career pathways, enabling reskilling and upskilling, and cataloguing emerging technology skills.
- Almost all APEC economies collect survey data regularly and have done so for decades, allowing for long-term trend analysis.
- Reliance on currently available skill data collected through surveys limits understanding of shorter-term changes in skill requirements, especially digital skills.
- Public–private partnerships can provide a cost-effective way for APEC economies to introduce big data analysis into their strategic processes.

BIG LABOR MARKET DATA

Big data aggregated from real-time labor market information complements traditional labor market information (LMI). Traditional LMI is drawn from sources that are reliable and representative. However, these sources often lack detail and tend to be produced infrequently. Real-time LMI, in contrast, can be deployed constantly resulting in monthly or even daily data on labor market trends. Additionally, real-time LMI techniques can generate detailed data about job openings and worker skills. However, real-time LMI often captures a partial picture of the labor market because more highly skilled, formal jobs are more likely to be found online. Thus, real-time and traditional LMI are both necessary and in fact strengthen each other: the frequency and depth of real-time LMI complements the breadth (representativeness) of traditional LMI.

A particularly relevant form of real-time labor market data is analysis of online job postings. These datasets are developed by collecting job vacancies from the web. Extracted information includes:

- **Geography** or area where the job advertised is located
- **Education requirement**
- **Experience required**
- **Competencies required**
- **Salary information**

Comparison of traditional labor market data and big data

Type of data	Years of data	Ease of time series analyses	Data representativeness	Compatibility across economies	Real-time data access	Regular taxonomy (classification) updates	Data granularity
Traditional labor market data	~50	High	Apply statistical sampling methods and weights	✓	×	×	Low
Big data	~10	Medium	Captures digitized labor market; can benchmark against public data to gain insight	✓	✓	✓	High

POLICYMAKERS USING BIG LABOR MARKET DATA

There are various examples of policymakers using big labor market data to increase granularity, timeliness and relevance of labor market information. Primary applications of big data that may be of particular use to policymakers include addressing time lags in traditional data, responding quickly to economic shocks such as the COVID-19 pandemic, reskilling and upskilling workers, and matching job seekers to critical occupations.

Utilizing real-time insights for quick policymaking: National Skills Commission of Australia

The Australian government uses big data, including near real-time information from online job advertisements, and has worked with labor market analytics firm EBG in many capacities. For example, the National Skills Commission of Australia**:**

- Uses big labor market data in their Jobs and Education Data Infrastructure (JEDI) tool. JEDI provides real-time data on the Australian labor market, helping policymakers to navigate the changing economy by providing data on the labor market, workforce changes, and current and emerging skills needs.
- Developed Nowcast of Employment by Region and Occupation (NERO) which directly addresses the issue of time lag in public data. Employment data in 355 occupations across 88 regions are only available every five years from the Australian Bureau of Statistics Census of Population and Housing, whereas this tool enables the user to understand monthly employment data.
- Used a mix of traditional sources of data and job posting data to identify 25 emerging occupations within the Australian labor market that are not well articulated in standard occupational taxonomies. The need to learn new skills was expedited during the COVID-19 pandemic. The emerging occupations analysis helps ensure that people are equipped with the right tools and skills for jobs of the future, and contributes to building a skilled, resilient and adaptable workforce. Using big data to complement traditional sources of data, the government could quickly understand changes in the labor market as well as broader labor market trends, and identify growing occupations into which they could funnel workers.
- Released the Australian Skills Classification (ASC) as a beta product, using job posting data to help validate the currency of skills in the Australian labor market. The ASC complements the Australian and New Zealand Standard Classification of Occupations (ANZSCO), providing a new level of detail about the skills that underpin Australian jobs. Intended to be a ‘common language’ for skills, the ASC provides ways to explore the connections, and transferability of skills, between occupations. The ASC will continue to be expanded and improved; and in conjunction with big data and other information sources, it will help identify changes within jobs, and new and emerging jobs, more readily.

Digital and data policymakers

- The **US Bureau of Economic Analysis** is developing an approach to using labor market data to measure the value of the data economy. Their approach uses job posting data from EBG to select data-related occupations based on skills and tasks; and they then estimate labor costs of these occupations using Occupational and Employment Statistics (OES) data.
- The **Measurement and Analysis of the Digital Economy (MADE) Group at the Organisation for Economic Co-operation and Development (OECD)** has considered using big labor market data to understand technology adoption and measure the effect of 'free' digital products and services on welfare, which is largely not reflected in national accounts. There are limitations posed by survey data around technology adoption especially with open source software. Job posting and social profile data can help to measure the adoption of open source technologies such as Python, TensorFlow and Perl, and the changes in demand over time as a share of all information technology (IT) jobs. Using a salary prediction model based on these technologies, value-added can be calculated.

Macroeconomic policy analyses

- Economists at the **Massachusetts Institute of Technology** used job posting data to understand wage rigidity. They found that posted wages change infrequently, wages for the typical job remained unchanged for 20 quarters, and that posted wages were especially unlikely to fall within a given job, implying downwards rigidity in the posted wage. They also found that posted wages were nearly acyclical for the typical job, suggesting substantial rigidity in the wage for new hires at the job level.
- Big labor market data can be used to understand the effects of wage transparency, such as with Colorado's new Pay Transparency Law. This law requires employers to (1) post compensation and benefits information for each job posting for Colorado jobs and (2) internally post promotional opportunities to current Colorado employees on the same day and sufficiently in advance of promotion decisions. Changes in wages based on these policy changes show up immediately in online job posting data, which enable the changes in wages or other job characteristics to be tracked.

Reskilling workers

- The former **Australian Department of Employment, Skills, Small and Family Business (**now the **Department of Education, Skills and Employment)** partnered with the Boston Consulting Group and EBG to develop a report on reskilling Australia. The Australian government decided to use big labor market data to supplement the commonly used public skills data, O*NET, because the EBG data provide information that O*NET cannot, such as occupations coded to the local taxonomy (ANZSCO), education requirements coded to Australian degree systems (such as vocational education and training (VET), bachelor's degree or above) and a dynamic and expansive skills taxonomy that include emerging skills like data analytics and blockchain.

- The **Australian Department of Education, Skills and Employment** also publishes weekly information on jobs in demand by location through the Jobs Hub app. This app was developed in response to the COVID-19 pandemic and the need to help people transition quickly to new roles during the labor market disruption. It allowed people to understand changes in their local labor market and how their existing skills could be applied to new roles.
- In **New Zealand,** Tokona Te Raki is using big labor market data to drive longer-term systemic change to boost Māori success and tackle inequality. They aim to produce an understanding of current labor market data to support better employment outcomes for Māori and inform the business case for further investment in tools to enable future indigenous workforce development.

Understanding emerging technologies and technological change

- The **Australian Department of Industry**, in conjunction with their agency AustCyber, uses big labor market data in a tool called CyberSeek which enables data exploration of the cybersecurity skills gap. The tool helps policymakers to answer questions about costs and ease of hiring in their region, and whether they should source workers from other regions.
- The **Singapore Smart Nation and Digital Government Office (SNDGO)** worked with LinkedIn Economic Graph as part of Singapore's Artificial Intelligence (AI) Strategy. The goal of the collaboration was to explore where the necessary AI talent works and what skills they have across regions. Big labor market data enabled understanding of emerging skills and trends that is otherwise not possible with traditional data. In addition, the Infocomm Media Development Authority of Singapore has been using EBG data since 2017 for early-stage analysis to help glean insights into local-talent demand to facilitate the making of policy decisions to help career counselors provide guidance around media and information and communications technology (ICT) roles.

Matching job seekers to employers and defining talent gaps and critical occupations

- The **European Union (EU)** partners with labor analytics company EBG to develop Cedefop's online job vacancy analysis system. Cedefop, an EU agency focused on the development of European VET policies, has been working with big labor market data since 2015. Ultimately, the end goal is to create a pan-European online job vacancy collection and analysis system that will provide data on Europe's current and emerging skill needs and information relevant to policymaking. In partnership with EBG Europe, they conducted a feasibility study; identified main job vacancy providers across member economies; created a methodological approach to crawling, fetching, and scraping; and collected job vacancy data.

- The **Australian Government** developed a platform called Your Career that helps people find career information and publishes employment and education data. The website allows users to look at occupations with current job vacancies based on a series of user input preferences.
- The World Bank and the government of **Malaysia** released a report on Malaysia's skill shortages and critical occupations. They used postings data to learn the skill and experience requirements of high-demand occupations and help create their list of critical occupations. This list is then used to align workforce development policies with employer demands.
- **Indonesia** also created a critical occupation list to highlight shortages and potentially strategic investment areas. Additionally, they have an Online Skills and Vacancy Outlook initiative that collects online job postings by occupation. These two initiatives work to analyze skill imbalances and help policymakers to make investments in training programs and adjust incentives.

RECOMMENDATIONS

For governments and policymakers looking to use big data to expand their understanding of the labor market, there are a few main challenges to consider and associated recommendations:

Challenges	Recommendations
High up-front costs to collect online labor market data, including: • Development of technical capabilities to automatically scrape data and identify online text as job postings, résumés, etc. • Storage space for storing large raw text data • Finding the best sources of job postings or other online labor market data, including job boards and individual employer website	• Work with a third-party intermediary familiar with data collection in other geographies to build aggregation system for job postings • Use local knowledge of labor market to identify high-density sources of labor market data, such as the economy's primary job board
Cleaning, deduplicating and preparing raw text data for analysis requires advanced modeling expertise including: • Technical capability to see a job posting or social profile/résumé on multiple websites and deduplicate across sources • Ability to read raw text in the source language and categorize the data into labeled fields, such as employer, occupation, location, skills, etc.	• Apply a parser from a third-party intermediary built on at least one local language as a pilot case • Limit sources to one large job board or one set of employers to ensure the ability to parse postings can be readily standardized • Work with local language experts or translators to translate some components of a parser into a local language

Challenges	Recommendations
Merging data with current government sources and taxonomies and analyzing big labor market data • Big labor market data may not be representative of the population and may be limited to more urban, digital or high-skill jobs. This may be a particular challenge for developing economies. • Maintaining a database of big labor market data requires near-constant updating, and ever-growing storage capacity	• Utilize crosswalks and parsers that parse to both local taxonomies and international standard taxonomies, which third-party intermediaries often have • Match the distribution of big labor market data to public datasets to understand gaps and differences • Plan for expanding storage costs as improvements in technology mean that an additional number of job postings/résumés are collected monthly or annually
Visualizing data and creating models and tools on top of big labor market data to answer research and policy questions • Dashboards and interfaces make real-time data accessible to a wider range of agencies rather than just technical experts, but need to be built and maintained • Big data analysis requires a different understanding of data limitations, such as changes in data collection over time affecting the time series analysis	• Build a tool in partnership with a data visualization organization to easily view different cuts of data, such as top-growing skills or top industries

In general, regardless of the type of labor market data being utilized we recommend the following as immediate next steps for policymakers:

1. **Understand and assess the full scale of costs.** In any of the partnership modalities (with academia or non-governmental organisations, for example), there are up-front costs as well as dynamic costs. It is important to consider the costs of initial creation and set-up, people and analytical resources, and maintenance of taxonomical and data updates.

2. **Begin with a small-scale pilot project**. This pilot project should aim to solve a very specific problem, like determining top skills required for each occupation in the economy. Initially, the government may face some hesitation from stakeholders used to working exclusively with traditional data. It is important that the first project starts small to gain trust in the data.

3. **Once trust is gained, start a larger project.** After the initial data scoping to ensure big labor market data can be helpful and used to solve a problem, the government can consider larger scale projects. Further projects can include work like the examples described earlier, including identifying labor shortages, emerging skills or growing occupations.

1. INTRODUCTION AND PROJECT OVERVIEW

The COVID-19 pandemic has drastically changed the labor market and has exacerbated megadrivers of change that were already on the rise. Rapid technological change characterized by digitalization and automation of jobs has become even more pronounced. Amid this economic crisis, access to real-time labor market data has become even more pivotal in shaping policy.

The importance of real-time data in addressing the COVID-19 economic crisis cannot be overstated. Almost all government sources of data are released with a lag of several months or sometimes years. In contrast, many real-time labor market data sources include only a one- or two-day lag and update constantly.

In addition to managing economic crises, governments and policymakers are also tasked with anticipating future trends. As discussed in the 2021 APEC Economic Policy Report (AEPR) on Structural Reform and the Future of Work, the increase in importance of digital skills and digital jobs, and the pace at which the skills required are changing, has accelerated change in the labor market. Relentless reskilling and upskilling will continue to be the norm as jobs and workplace relationships transform.[1] Standard information provided by labor market surveys does not capture this rapid change and is slow to respond in the face of these trends.

Enhancing APEC economies' abilities to access and understand labor market information is crucial to improving labor market outcomes for workers in those economies. In this project, 'Filling APEC's Data Gaps to Address Future of Work Challenges', Emsi Burning Glass (EBG) has conducted a mixed methods study of existing labor market data sources, and integration of new data into labor market systems, and assessed opportunities for public–private partnerships. The study is based around a review of the literature and specific case studies of innovative labor market data integration and also draws from quantitative analysis of a range of real-time labor market data.

The report includes:

- A description of traditional sources of labor market data across all 21 APEC economies, including survey methodologies, and key strengths and challenges of these data
- A description of alternate sources of labor market data, including online job posting data, human capital management data and other sources relevant to labor analytics
- Best-practices associated with using real-time data and narrowing adoption time for new users
- Sample use cases for real-time data across a variety of industries and applications
- Collaboration methods and associated costs to help support integration and public–private partnerships for the adoption of big data

[1] APEC Economic Committee, "APEC Economic Policy Report 2021: Structural Reform and the Future of Work" (Singapore, APEC, November 2021), https://www.apec.org/publications/2021/11/2021-apec-economic-policy-report.

Key questions addressed in this study include:

- What gaps exist in current labor market information in APEC economies particularly with regard to digital jobs and skills?
- What new data sources are available, such as big data and digital labor market platforms?
- How can these new data sources be integrated into current labor market systems?
- What opportunities for public–private partnerships exist and how have they succeeded in the past?

2. CURRENT LABOR MARKET INFORMATION SOURCES

This section draws from literature and information on currently available labor market data, particularly with regard to jobs and skills in the digital economy. There are three primary sources that are reviewed in this section, making up the collection of frequently used labor market information: labor market survey data, competency data and emerging technology surveys.

First, we explore publicly available data contained in labor market information (LMI) systems across all 21 APEC economies. Next, we unpack labor force survey methodologies in five economies in order to compare different approaches.

In addition to basic traditional labor market data, this report also considers other primary data sources currently used to understand jobs and skills. This includes a description of skill and competency datasets as well as artificial intelligence and emerging technology data that is collected primarily by private entities through ad hoc surveys, but is also used by government agencies to understand future of work challenges.

TRADITIONALLY COLLECTED DATA ON EMPLOYMENT AND UNEMPLOYMENT

The first data source used by stakeholders trying to understand labor market dynamics is government collected data usually from labor force surveys (LFS). These surveys include information on the volume of people working in a specific occupation. This data is usually maintained by individual government organizations as well as some international organizations. These types of data sources answer research questions like: *What are the jobs in my labor market? Which occupations have seen decline or have increased in the long term?*

Coverage and frequency

Table 2.1 shows basic information about the availability of data from traditional LFS across all APEC economies.

Table 2.1 Coverage and frequency of updates

APEC member economy	Coverage				Frequency of updates
	Collection start year	Location/ geography	Industry breakdown	Occupation breakdown	Time series data
Australia	1960	✓	✓	✓	Monthly
Brunei Darussalam	2014	×	✓	✓	On ad hoc basis (most recent is 2019)
Canada	1945	✓	✓	✓	Monthly
Chile	1966	✓	✓	✓	Quarterly
People's Republic of China	1996	✓	✓	✓	Annually
Hong Kong, China	1975	×	✓	✓	Monthly
Indonesia	1986	✓	✓	✓	Annually
Japan	1953	✓	✓	✓	Monthly
Republic of Korea	1963	✓	✓	✓	Monthly
Malaysia	1962	×	✓	✓	Monthly
Mexico	1972	✓	✓	✓	Monthly
New Zealand	1985	✓	✓	✓	Quarterly
Papua New Guinea	Does not conduct a regular LFS	×	×	×	Not conducted regularly
Peru	1996	✓	✓	✓	Annually
The Philippines	1956	✓	✓	✓	Quarterly
Russia	1992	✓	✓	✓	Quarterly
Singapore	1974	×	✓	✓	Quarterly
Chinese Taipei	1962	×	✓	✓	Monthly
Thailand	1963	✓	✓	✓	Quarterly
United States	1940	✓	✓	✓	Monthly
Viet Nam	2007	✓	✓	✓	Quarterly

LFS=labor force survey

Sources: See Appendix for a complete list of sources for all data contained in this table.

All economies except for Papua New Guinea conduct a regular LFS, with Brunei Darussalam being the most recent to begin collecting data (in 2014) and the United States starting earliest (1940). Every economy that does conduct a LFS regularly collects data in such a way that it can be broken down by industry and by occupation. Of the 20 APEC economies with a regular LFS, nine conduct the survey on a monthly basis, seven on a quarterly basis, two annually, and one relies on ad hoc survey data. Six economies do not have regional breakdowns available.

Taxonomies, microdata, and crosscutting availability

Table 2.2 contains more detailed information on taxonomy availability, microdata and the ability to crosscut the data by multiple variables. Most economies include data with the ability to be broken down by region. Many economies also maintain occupation taxonomies that can be linked to international standards, such as ISCO-08. Most have the ability to crosscut the data by at least location, industry and occupation ('Medium' in the table) and many by additional variables like age, sex and educational attainment ('High' in the table). Access to the microdata varies across economies; many have application processes, including payments, for access, with only a few making the data fully publicly available. A handful of economies have only limited access to the microdata.

Table 2.2 Detailed taxonomy information and access

APEC member economy	Most granular taxonomy available					Microdata availability	Ability to crosscut the data by variables[a]
	Location/ Geography	Industry		Occupations			
		Taxonomy	ISIC Rev. 4 (intl.) compatibility	Taxonomy	ISCO-08 (intl.) compatibility		
Australia	State, territory, greater capital city	ANZSIC subdivision	Broadly aligns with ISIC; crosswalk also available	ANZSCO occupation major group	Major groups broadly like ANZSCO; crosswalk also available	Apply to access, free and charged options available	High
Brunei Darussalam	N/A	BDSIC 2011	Based on ISIC Rev. 4	BDSOC 2011	N/A	N/A	Low
Canada	Province and territory	NAICS 2012 2	Crosswalk available	NOC 2016 3, NOC 2021 v0 just released	Crosswalk available	Publicly available	High
Chile	Region	CAENES, adaptation of the CIIU4	Data available	CIUO-08	Data available	Apply for access	Medium

People's Republic of China	Region	GB/T 4754-2011	Based on ISIC Rev. 4	ISCO-08	Data available	Limited access	Medium
Hong Kong, China	N/A	HSIC Version 2.0	Based on ISIC Rev. 4 with local adaptation; crosswalk also available	Occupation index of Hong Kong Population Census	Based on ISCO-08 with local adaptation	Apply to access with possible charge	High
Indonesia	Province	KBLI, 3-digit	Links to ISIC Rev. 4	KBJI, which referred to ISCO-88, 3-digit	Links to ISCO-88	Available for purchase	Medium
Japan	Region	JSIC	Crosswalk available	JSOC	Crosswalk available	Limited access	Medium
Republic of Korea	City and province	KSIC 1992 2-digit	Crosswalk available	KSOC 1993 2-digit	Crosswalk available	Apply to access	High
Malaysia	N/A	Malaysian Standard Industrial Classification	Very similar to ISIC Rev. 4	Malaysian Standard Occupational Classification	Very similar to ISCO-08	Request access with possible charge	Low
Mexico	Entity	NAICS 2 digit	Crosswalk available	Mexican Classification of Occupations major groups	Crosswalk available	Publicly available	High
New Zealand	Region	ANZSIC Subdivision	Broadly aligns with ISIC; crosswalk also available	ANZSCO occupation major group	Major groups broadly like ANZSCO; crosswalk also available	Apply to access	High
Papua New Guinea	N/A	N/A	N/A	N/A	N/A	N/A	Low
Peru	Ubigeo coding (admin. subdivisions)	ISIC Rev. 3	Data available	ISCO-88 3-digit	Data available	Publicly available	Medium
The Philippines	Region, Province, City	1994 Philippine Standard Industrial Classification (PSIC) major divisions	Based on ISIC Rev. 4 with local adaptation	Philippine Standard Occupational Classification major groups	Based on ISCO-08 with local adaptation	N/A	High

Russia	Territory	ISIC-Rev.3 compatible classification	Data available	Russian ISCO-88 compatible classification	Data available	Limited access	High
Singapore	N/A	Singapore Standard Industrial Classification 2020, adopts ISIC Rev. 4, 1 digit	Broadly aligns with ISIC Rev. 4; Crosswalk available	Singapore Standard Occupational Classification, 2020	Broadly aligns with ISCO-08; Crosswalk available	N/A	High
Chinese Taipei	N/A	Statistical Classification of Industry Section	Based on ISIC	Standard Occupational Classification System	N/A	N/A	High
Thailand	Region and province	Thailand Standard Industrial Classification (TSIC)	ISIC Rev. 4	ISCO-88	ISCO-88	Publicly available	High
United States	State, metro	NAICS 2-6 digit codes	Crosswalk available	SOC detailed occupations	Crosswalk available	Apply to access	High
Viet Nam	Region and province	2007 Vietnam Standard Industry System 1 digit	Based on ISIC Rev. 3	Vietnam Standard Classification of Occupations 1 digit	Based on ISCO-88	N/A	Medium

[a] Low: Unable to crosscut by location, industry and occupation; Medium: Able to crosscut by location, industry and occupation but no other variables; High: Able to crosscut data by location, industry and occupation, as well as additional variables

Sources: See Appendix for a complete list of sources for all data contained in this table.

Survey methodologies in five economies

Table 2.3 Survey methodology comparison

Economy	Sample size	Scope	Weighting and adjustments
Australia	26,000 households (50,000 people or 0.32% of the population)	Civilian population 15 years+, excluding: - Members of the permanent defense forces - Certain diplomatic personnel of overseas government - Overseas residents Multi-stage area sample of: - Private homes - Discrete Aboriginal and Torres Strait Islander communities - Non-private dwellings (hotels, motels, hospitals, retirement communities)	- Combines data from previous 6 months with current month's data to get current estimates - Weighting to align with current-month population benchmarks
Chile	54,000 households	People ages 15 years+ who live in households in private homes	Adjusted for unknown eligibility, ineligibility, non-response and population groups
Korea	35,000 households (0.2% of all households)	People ages 15 years+	Number of total employed for economy is estimated using survey results and sample weights by sex, age and region Updated every month based on monthly population projections
The Philippines	51,000 households	People ages 15 years+ excluding overseas workers	Weigh adjustment factors for non-interview households and population projection
Thailand	83,880 households	People ages 15 years+; private households and special households including group living, dormitory, factory compounds	Adjustments for non-response and projected population

Sources: See Appendix for a complete list of sources for all data contained in this table.

In Table 2.3Table 2.3, five APEC economies – Australia; Chile; Korea; the Philippines; and Thailand – are reviewed in more detail to understand the scope of their LFS process. The economies were selected based on their range of population sizes and to explain the data collection methods underlying their LFS. In general, each economy approaches their LFS similarly. They include population ages 15 years and above, largely based on the civilian non-overseas population. Households included range from only private households (Chile), to additionally those in special households or non-private dwellings, such as hospitals, dormitories, retirement communities and hotels. Each economy studied has its own sampling methodology tailored to the distribution of the population across regions and in urban v. rural areas. The survey methodologies include face-to-face interviews and phone interviews; and the surveys are performed by individual enumerators as well as computer-assisted technologies. All five economies adjust their data to align with current population benchmarks.

STRENGTHS AND LIMITATIONS OF TRADITIONAL LFS DATA

Traditional labor market data, including government-collected LFS data, skills and competency data and technology adoption survey data, provide rich understanding of the labor market at an aggregate level. This section outlines some of the primary strengths and limitations associated with these data.

Strengths

These datasets have some features that are particularly valuable to users:

Representativeness: As seen in the survey methodology section, governments working to collect data through LFS use statistical sampling methods to ensure that the data is representative across demographics and regions. These techniques include random sampling, and coverage that considers rural v. urban areas, a wide range of living situations, and other variables that could vary across the economy, such as number of overseas residents. The survey methodologies also compare to other population surveys and use weights to account for representativeness. Taken together, these approaches mean that the data collected accounts for all corners of the labor market.

Length and stability over time**:** One of the key features of using labor market data is the ability to track how an occupation or industry has changed over time. The taxonomies used in LFS are created to be compatible across time. This means that, even if updates are made, ample data documentation is provided to allow for the continuation of time series data and to attempt to compare apples-to-apples over many years. Additionally, many economies have data dating back at least 50 years, and some even longer. The length of this time series allows for a deep understanding of historic labor market trends.

Accessibility**:** Aggregate data on labor market outcomes, including employment by location, industry and occupation are widely available for free to the public online. Some economies additionally make their microdata accessible to researchers, allowing for more granular exploration of the data.

Compatibility**:** While many governments use their own proprietary taxonomies for industry and occupation data, most also provide some compatibility with international standards, such as the International Standard Classification of Occupations (ISCO) and International Standard Industrial Classification (ISIC). This allows for cross-economy comparison.

Limitations

Traditional sources of labor market data also have some limitations:

***Slow to collect and update data content*:** The frequency with which data are collected and made available are limited by the labor-intensive method of survey data collection. Most economies update the data either monthly or quarterly, but some update only annually or on an ad hoc basis. Even in the economies that update monthly or quarterly, some indicators are updated less frequently (annually or ad hoc). For some economies, there is also a lag in the availability of data so even when new data is released, it is not from the most recent months. This limits the analyses that could be performed to retroactive views of the state of the economy.

***Slow to update taxonomies*:** In general, due to prioritizing time series analysis and the ability to maintain consistency, government taxonomies are updated infrequently. This means that emerging occupations or nascent industries are difficult to capture, especially in times of labor market shocks.

***Level of granularity*:** Most economies allow for the data to be crosscut across industry, occupation and region. However, these sources of data lack granularity to understand variables such as skills at the occupation level, education, experience and other labor market variables.

SKILLS, KNOWLEDGE, ABILITIES AND COMPETENCY DATA

Various sources offer data on skills, knowledge, abilities and competencies. These datasets are helpful for determining which skills are important to different occupations and understanding how occupations' skill requirements compare to each other. These datasets address research questions such as: *What skills are required for this occupation? How do occupations compare to each other in terms of skill requirements?*

Altamirano and Armaral (2020) note that O*NET and ESCO are the two most well-known skill taxonomies. These taxonomies classify occupations and skills and map them to each other. In Table 2.4, these taxonomies are compared by coverage and accessibility.

Table 2.4 O*NET and ESCO skill taxonomies

Skills database	Year of first full data release	Occupation count	Skill/competency count[a]	Number of languages available
O*NET	2008	1,016	277	2
ESCO	2017	2,942	13,485	27

ESCO=European Skills, Competencies, Qualifications and Occupations; O*NET=Occupational Information Network
[a] For O*NET, this includes including abilities, skills, work activities, context and styles, tasks, knowledge, interests and values.
Sources: European Commission (2017); O*NET (n.d.-b).

Occupational Information Network (O*NET)

O*NET is an occupational and skills framework developed by the US. Mariani (1999) notes that O*NET replaced the Dictionary of Occupational Titles as a source of occupational information. It was developed in the mid-1990s and became available for public use in 1998, with the first full version released in 2008.

Dickerson et al. (2012) describes O*NET as comprising professional assessments and employee self-reported assessments. This two-stage design is further expanded in O*NET's Data Collection Overview: the businesses are randomly selected from those expected to employ people in the target occupations, and the workers are randomly selected from the businesses. To make the survey manageable for respondents, the questions are split into three surveys, each with different questions. The selected sample is then split and randomly assigned to one of the three surveys. In cases where the occupations have small samples or are otherwise difficult to survey, occupational experts complete surveys. The O*NET database is updated annually as new data are collected.

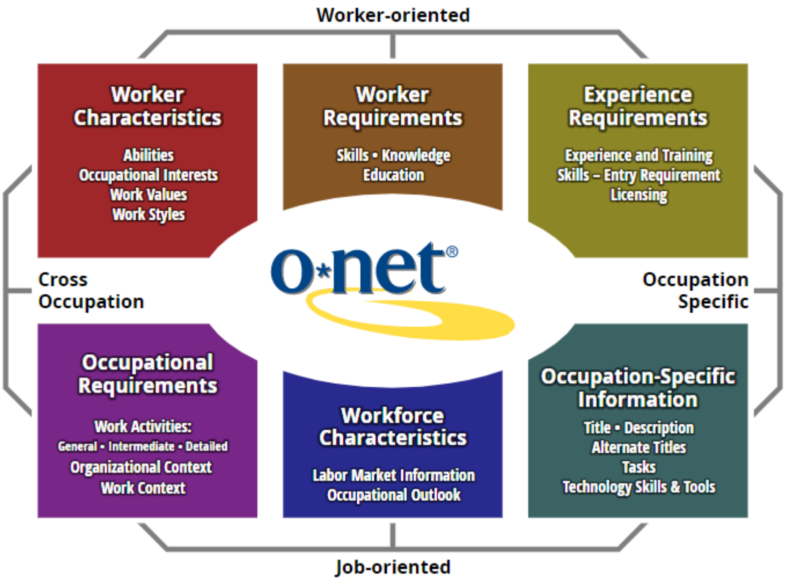
This data is recorded for 1,016 occupations and the survey spans 277 dimensions (including abilities, skills, work activities, context and styles, tasks, knowledge, etc.). Two out of three of the dimensions include the level and intensity of use. These dimensions span across six categories described in O*NET's Content Model. Developed by organizational analysts, the O*NET Content Model is the framework of O*NET (Figure 2.1).

O*NET categorizes skills by six broad groups, including Basic Skills, Technical Skills and Social Skills. Each skill can be examined by importance to occupation. O*NET also contains data on abilities, including Physical and Cognitive Abilities, that can be analyzed by occupation. A sample of O*NET skills and abilities data for Sales Representatives is in Figure 2.2.

The importance of specific skills and abilities and the level of knowledge required are generated from a questionnaire (sample in Figure 2.3). Occupational analysts then use the responses to generate the occupation values.

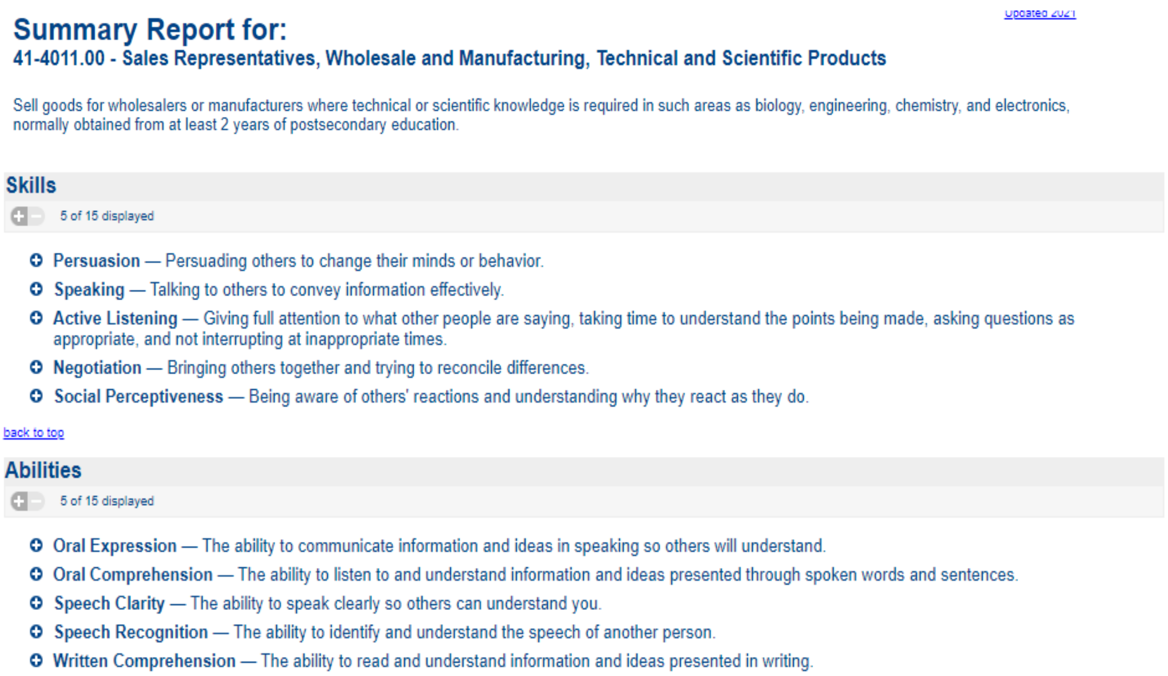
The Organisation for Economic Co-operation and Development (OECD) uses O*NET skill and ability definitions to power their Skills for Jobs tool. One feature of this tool allows users to input their current occupation and an occupation they would like to work in, and it outputs which skills may need to be strengthened to move between occupations.

Figure 2.1 O*NET database (‘The O*NET® Content Model’)



Source: O*NET (n.d.-b).

Figure 2.2 O*NET example description (‘Summary report for 41-4011.00’)


Updated 2021

Summary Report for:

41-4011.00 - Sales Representatives, Wholesale and Manufacturing, Technical and Scientific Products

Sell goods for wholesalers or manufacturers where technical or scientific knowledge is required in such areas as biology, engineering, chemistry, and electronics, normally obtained from at least 2 years of postsecondary education.

Skills

5 of 15 displayed

- Persuasion — Persuading others to change their minds or behavior.
- Speaking — Talking to others to convey information effectively.
- Active Listening — Giving full attention to what other people are saying, taking time to understand the points being made, asking questions as appropriate, and not interrupting at inappropriate times.
- Negotiation — Bringing others together and trying to reconcile differences.
- Social Perceptiveness — Being aware of others' reactions and understanding why they react as they do.

back to top

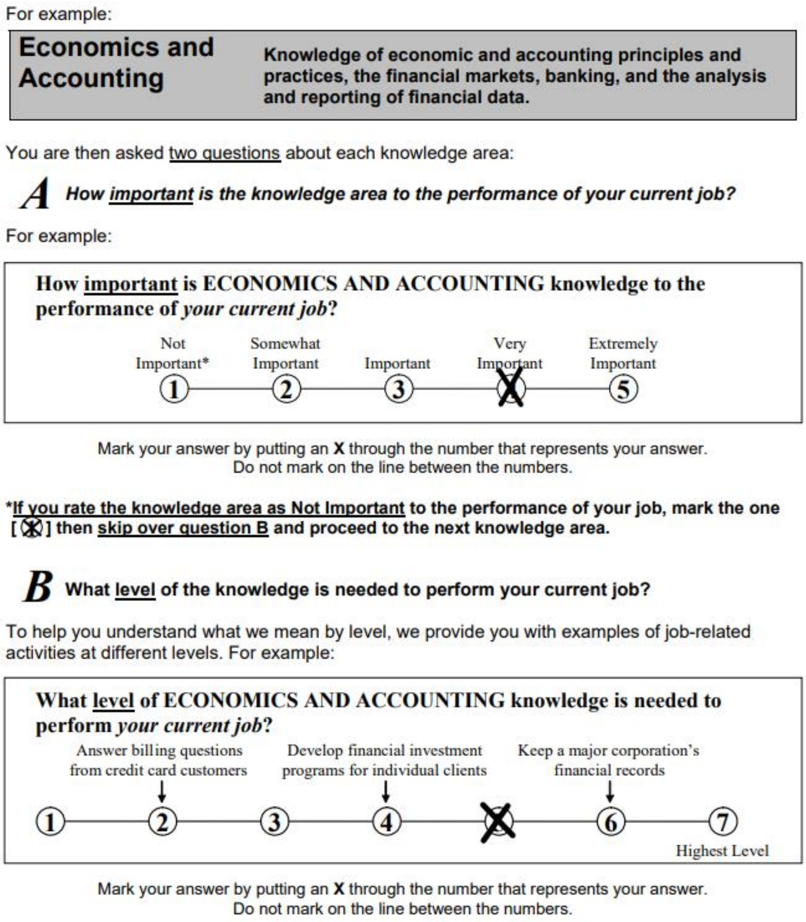
Abilities

5 of 15 displayed

- Oral Expression — The ability to communicate information and ideas in speaking so others will understand.
- Oral Comprehension — The ability to listen to and understand information and ideas presented through spoken words and sentences.
- Speech Clarity — The ability to speak clearly so others can understand you.
- Speech Recognition — The ability to identify and understand the speech of another person.
- Written Comprehension — The ability to read and understand information and ideas presented in writing.

Source: O*NET (n.d.-a).

Figure 2.3 O*NET sample questionnaires


For example:

Economics and Accounting	Knowledge of economic and accounting principles and practices, the financial markets, banking, and the analysis and reporting of financial data.

You are then asked two questions about each knowledge area:

A ***How important is the knowledge area to the performance of your current job?***

For example:

How important is ECONOMICS AND ACCOUNTING knowledge to the performance of *your current job*?

Not Important* ① Somewhat Important ② Important ③ Very Important ④ Extremely Important ⑤

Mark your answer by putting an **X** through the number that represents your answer.
Do not mark on the line between the numbers.

***If you rate the knowledge area as Not Important to the performance of your job, mark the one [①] then skip over question B and proceed to the next knowledge area.**

B **What level of the knowledge is needed to perform your current job?**

To help you understand what we mean by level, we provide you with examples of job-related activities at different levels. For example:

What level of ECONOMICS AND ACCOUNTING knowledge is needed to perform *your current job*?

Answer billing questions from credit card customers (②) — Develop financial investment programs for individual clients (④) — Keep a major corporation's financial records (⑥)

① ② ③ ④ ⑤ ⑥ ⑦ Highest Level

Mark your answer by putting an **X** through the number that represents your answer.
Do not mark on the line between the numbers.

Source: O*NET (2017).

Box 2.1 Technology adoption survey data

The technology adoption survey data seek to measure emerging technology adoption across companies and locations. Recent examples include measurement of the use of cloud technologies, artificial intelligence (AI) technologies and other technologies associated with increases in efficiency and productivity as well as the potential for automation. Stakeholders use these data to understand how to prepare for changes in digital needs in specific occupations and the labor market in general.

Technology adoption data can answer questions such as: *How have new technologies such as AI penetrated the workforce overall? What industries have adopted the use of technologies such as robotics and cloud technologies?*

In general, technology adoption survey data are collected by private sector stakeholders. These companies are interested in technology adoption for thought leadership, consulting, investment and other purposes and often make the data available to the public and to government leaders. Responses are gathered primarily from company participants representing their industry, organization or specific function within their organization. Recent examples of such survey data include:

McKinsey Global Survey on the state of AI in 2020 (Balakrishnan et al., 2020): This survey was fielded online and included 2,395 participants representing a range of regions, industries and company sizes. Respondents were asked about their company's adoption of AI within functions and their organization's AI use in general.

International Data Corporation (IDC) survey (Jyoti and Shirer, 2020)**:** IDC surveyed more than 2,000 IT and business line decision makers about adoption of AI. The survey was designed to understand what drives customer buying behavior for AI on a worldwide basis across different IT and business line personas, including data scientists, data architects, data engineers, and operations.

O'Reilly's 2021 AI Adoption in the Enterprise survey (Loukides, 2021)**:** This surveyed more than 2,500 business leaders and aimed to assess challenges in AI. It found that a lack of skilled people and difficulty of hiring topped the challenges in AI.

KPMG AI Adoption survey (KPMG, 2021)**:** KPMG surveyed 950 people in 2021 in tech, retail, financial services, industrial manufacturing, healthcare, life sciences and government. The goal was to assess the perceptions of AI technology and to target decision makers, including manager level or above employees.

Cognilytica AI Trends Forecast survey (Cognilytica, 2020)**:** This market research firm implemented a survey of public and private sector companies about their current stance on AI adoption and usage. This firm focuses on industry and adoption focused market research on AI and machine learning and has partnered with the Organisation for Economic Co-operation and Development (OECD) and other international organizations as a thought partner.

Deloitte's State of AI in the Enterprise survey (Ammanath et al., 2020)**:** Deloitte surveyed 2,737 IT and line-of-business executives across nine economies on AI adoption. Approximately 50 percent were IT executives, with the rest C-level executives. Respondents were required to meet one of the following criteria: determine AI technology spending and/or approve AI investments; develop AI technology strategies; manage or oversee AI technology implementation; serve as an AI technology subject-matter expert; or make or influence decisions around AI technology.

Annual Business Survey (Zolas et al., 2020)**:** The Census Bureau of the United States in partnership with the Center for Science and Engineering Statistics piloted this survey in 2017. The sample is ~300,000 employers in years 2018–2021, making it one of the largest technology adoption datasets. Employer response is legally required, which avoids the sample biases and low response rates that many private surveys face[1].

While not representative of the complete collection of surveys on AI adoption, these examples serve to illustrate the main methodological approach. Most surveys focus on IT leaders and other high-level executives and are conducted through online surveys of fewer than 3,000 people. The information provided is about high-level strategies but does not necessarily go down to specific AI skills requirements, nor are these requirements able to be assessed by occupation, industry and geography.

Other data sources that are important to understanding the technology and AI landscape are Capital IQ, Crunchbase and Quid (three companies that make investment activity easily accessible). These sources are utilized to understand global investment in AI, or more specifically the degree and rate at which various economies are investing in AI. These sources can help economies adequately prepare workers for the future of work and for economies to remain competitive in a world where digital skills are increasingly augmenting productivity and thus boosting economic activity.

ESCO

O*NET also serves as the reference/basis for other taxonomies. The European Skills, Competencies, Qualifications and Occupations (ESCO) skill classifications is based on elements of O*NET and various other taxonomies. ESCO was first fully published in 2017 and was developed by a team of experts. These experts have substantial knowledge of each industry and are able to define the tasks and required skills of each occupation.

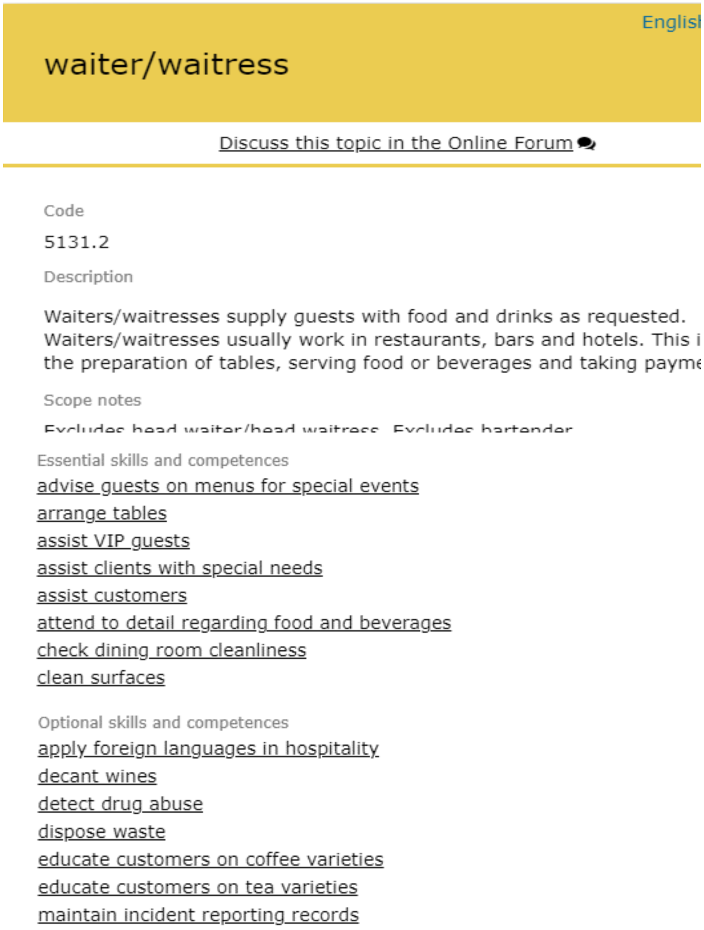
The ESCO Handbook includes more information about its development process: a core project expert team created the occupational profiles, working to ensure that the skills listed matched the constantly evolving skill requirements in the market. During the development process, the team consulted and analyzed many studies and already existing classifications. Additionally, various stakeholder groups reviewed the occupation and skill profiles and made suggestions for improvement. The extensive review and revision process ensures that the classifications are fitting and applicable across various sectors.

Altamirano and Armaral (2020) describe the difficulties with this approach, noting that developing and updating in this way demands a lot of resources and time. This system also requires lots of coordination and specialized knowledge, making updates difficult. Even so, ESCO is continuously updated through a process that includes collecting feedback, planning the release scope, and developing and quality testing.

ESCO creates an occupation and skill map that is applicable across the European Union. Altamirano and Armaral (2020) describe that the specific objectives include promoting labor mobility, making data transparent and accessible, and increasing data sharing across various sectors, including employers, educational institutions, and people seeking jobs. ESCO allows skill analysis by occupation, similar to O*NET. The occupations are based on the international ISCO-08 taxonomy with more detail, and each occupation lists essential as well as optional skills. ESCO skills definition includes knowledge, skills and competence. A sample of skills for a waiter/waitress is in Figure 2.4.

ESCO is available in 27 languages and linked to many other international taxonomies and classifications, making it very versatile and flexible.

Figure 2.4 ESCO sample skill and occupations profile



Source: ESCO (2017).

Strengths and limitations of skill and competency data

This section outlines the main strengths and limitations of skill and competency datasets.

Strengths

Level of granularity: Both ESCO and O*NET offer deep levels of granularity, with thousands of variables and unique skills collected for each occupation. This makes them extremely valuable for understanding in-depth the competencies, skills and abilities required for various occupations.

Compatibility: Both O*NET and ESCO are compatible with their respective government occupation taxonomies (and in the case of ESCO with the ISCO taxonomy). ESCO is also available in 27 languages, although this does not include many APEC economy languages.

Limitations

*Slow to update taxonomies***:** Both O*NET and ESCO rely on those employed in occupations as well as experts to craft their taxonomies. This extensive process means that both are very comprehensive, but slow to update. The slow update process makes it difficult to track how occupations change over time in terms of the skills they require, something that is becoming even more salient as digitalization transforms occupations.

*Limited relevance across economies***:** O*NET is based on the American economy, whereas ESCO is based on the European Union economy. Differences in industry cultures could affect the types of skills deemed important for a job, so these frameworks are limited in the economies they can be applied to.

3. IDENTIFYING NEW SOURCES OF LABOR MARKET DATA

This section explores available new big data sources on the labor market. The sources covered include online job ads (also known as vacancies or job postings data); employee work history data, such as social media profile or résumé data; human capital management data, such as applicant tracking information, payroll, and employee management data; and other sources from online and digitally available data.

ONLINE JOB ADS AND EMPLOYEE WORK HISTORY DATA

The jobs that employers post online and the résumés/curricula vitae (CVs) workers upload to the internet can provide rich insights into the landscape of an economy's labor market. Intermediaries extract a wealth of information from these sources, shown in Table 3.1.

Table 3.1 Information from job posts and résumés

Information extracted from job posts	Information extracted from résumés/CVs
• Employer name and industry/sector • Job title/occupation • Skill requirements • Education, certification and experience requirements • Compensation • Job location • Job type: Full-time, part-time, permanent, temporary, internship, remote, etc. • Duration of post	• Worker name (generally anonymized during processing and aggregation) • Residential location • Current and former job titles/occupations • Time spent in each role • Current and former employers and industries • Job locations • Competencies/skills • Education and alma mater

Several intermediaries compile and transform raw job posts and résumés into intelligible datasets in real time, as they are uploaded to the internet. When aggregated, these data inform the demand for occupations, skills and experience by employer, industry or location. The data can also be used to identify historical trends, make projections, and track how occupations, skills, industries, and geographies are changing. These data generally have trends that align with traditional sources. Clusters of talent with job-ready skills can be located and skill gaps identified. Regional assessments can be performed on the competitiveness for hiring particular occupations and talent. Wage analyses around occupations, skills, education and experience can be performed. Career pathways to secure well-paying jobs can be mapped for workers. Competitive analyses can be performed between economies, industries and companies. In Table 3.2 and Table 3.3 are a few vendors that compile these data and a description of how they differentiate themselves.

Table 3.2 Coverage of online job ads and employee work history data

Third party intermediary	Coverage		
	How to access data	Has global data? (economies available)	Benchmarked against public data
Indeed	By Request hiringlabinfo@indeed.com	60+	?
LinkedIn	Open Source https://www.linkedin.com/developers/	200+	×
Emsi Burning Glass (EBG)	By Request info@burning-glass.com	~30	✓
Glassdoor	By Request www.glassdoor.com/research/	~190	?

Table 3.3 Source of online job ads and employee work history data

Third party intermediary	Source of Data						
	Online job ads	Integrates traditional public data	Résumés	Social profiles	Employee sentiment surveys	Job seeker activity	College curriculums
Indeed	✓	×	✓	×	✓	✓	×
LinkedIn	✓	×	✓	✓	✓	✓	×
Emsi Burning Glass (EBG)	✓	✓	✓	✓	×	×	✓
Glassdoor	✓	×	×	×	✓	✓	×

Indeed

Indeed is a job search engine that aggregates online job ads from all over the web. It also allows employers to upload jobs, and workers to upload their résumés. The company claims 250 million unique monthly visitors, 175 million résumés, 320 million employer ratings and reviews, 10 jobs added per second, 750 million salaries, and websites in 60 markets.

Unique proprietary data:

- Job seekers' search activity can be used to understand the roles, industries and types of jobs being sought and where those jobs are.
- Job seekers' click data provide insight into the job posts and companies job seekers are likely to click on and apply to.

LinkedIn

LinkedIn is an online social network for professionals to connect with each other, find jobs and learn new skills. Recruiters also use the site to find talent. The company collects data from professionals who upload their work and education histories, and from employers that advertise jobs on their site. It claims to have 756 million members, more than 57 million companies, more than 15 million jobs, 38,000 skills, and 120,000 schools represented across more than 200 economies.

Unique proprietary data:

- User profiles tend to be updated with more frequency than résumés and may provide clearer insight into hiring trends, worker migration, the supply and skillset of talent within companies, industries and regions.
- Network analyses can be performed to understand professional networks and their impact on job seekers, employers and labor markets broadly.

EBG

EBG is a labor market analytics firm that specializes in the future of work and analyzing changes in labor markets across the globe. The company claims to have compiled over a billion historical and current job posts and pulls in millions of new posts daily from more than 50,000 sources around the world. Its database also includes over 100 million résumés and professional social profiles and traditional labor market and education data.

Unique proprietary data:

- Developed a robust proprietary taxonomy of almost 2,000 occupations and more than 17,000 skills that allow ease of tracking emerging, growing, stable and declining occupations and skills and allows for comparison across economies and against government taxonomies. Taxonomies are updated and maintained based on trends captured in job posting data.
- Job posting and talent data are structured so that they can easily be merged and analyzed alongside traditional labor market and education data.
- Demographic data can be leveraged to understand the gender and racial/ethnic makeup of companies, industries and occupations. Career pathway analyses can be performed to diversify homogenous roles and industries.

Glassdoor

Glassdoor is a job search engine and review site that allows workers to write reviews and rate current and former employers and jobs. The company collects data from job posts, employer sentiment surveys, and profiles created by companies. It claims to have 60 million unique monthly visitors, 1.5 million employers, and 90 million reviews, salaries and insights across 190 economies.

Unique proprietary data:

- It conducts sentiment surveys about work culture, benefits and attributes that provide insight into employer–employee relationships. These insights can help inform how worker happiness affects hiring, turnover and productivity. It can also be used for competitive analyses.
- Survey and salary data are also available by race/ethnicity and gender.

Limitations of online job ads and résumé data

While online job ads and work histories can serve as proxies for the demand and supply of talent, they do not reflect the full universe of a labor market. Certain industries, employers, workers and geographies are less likely to use the internet for employer–employee matching (e.g., construction, mining, agriculture), resulting in them being underrepresented in these datasets. Similarly, some may favor certain intermediaries over others. For example, LinkedIn data overrepresents the IT and finance industries; jobs for nurses, truck drivers and sale associates are overrepresented on online job boards; and some intermediaries serve niche markets. Also, a significant number of matches are still made via word of mouth, networks and help wanted signs. Estimates suggest 40 to 60 percent of jobs that do not require a college education are posted online, 30 to 40 percent of jobs that require some college or vocational training are posted, and 80 to 90 percent of jobs requiring college degrees are advertised online (Carnevale et al., 2014).

Moreover, online job ads do not always reflect the number of job openings. An employer may post a single job ad for several openings and companies sometimes use online ads for market research – these ads do not reflect actual openings at all. Therefore, it is important to understand how these shortcomings may present in the data, as well as how they may vary based on the intermediary.

HUMAN CAPITAL MANAGEMENT DATA

Human capital management (HCM) solutions providers capture a wealth of labor market information as a by-product of their business operations.

Information extracted from HCM operations

- Employer name, industry and number of employees on payroll
- Employee name and demographics
- Job title/occupation
- Job type: Full-time, part-time, permanent, temporary, internship, remote, etc.
- Payroll distribution including wages and tax payments
- Benefit packages
- Hours worked
- Hiring and turnover
- Employee start and end dates
- Time-off, vacation and sick days

HCM data can provide real-time insight into the employment activity of labor markets. The data can inform whether demand for talent or hiring is increasing, decreasing or stagnant by location, industry and employer. It can serve as an early indicator of how unemployment may change or has changed over a certain period. Wage analyses can inform whether a labor market is tight or loose based on wage inflation, deflation or stagnation. Insight into hours worked can inform whether overall economic activity is strong, waning or recovering. Compensation and benefits analyses can inform how incentives impact hiring, turnover and hours worked. These data can also highlight worker mobility, pay inequities, diversity, worker performance and the state of the gig economy. Table 3.4 lists a few vendors that compile these data and a description of how they differentiate themselves follows.

Table 3.4 Coverage and source of HCM data

Third party intermediary	Coverage			Source of data	
	How to access data	Benchmarked against public data	Has global data?	HCM	Employee surveys
Automatic Data Processing (ADP)	By Request ADP.Research.Institute@ADP.com	✓	✓	✓	✓
Kronos	By Request www.ukg.com/contact	?	✓	✓	✓
Homebase	By Request data@joinhomebase.com	?	US-only	✓	✓
ICIMS	By Request https://www.icims.com/company/contact-us/	?	✓	✓	✓

HCM=human capital management

ADP

Automatic Data Processing (ADP) processes payroll, time and attendance, benefits, taxes and other miscellaneous HCM data for more than 900,000 companies and some 100 million workers across 140 economies (ADP, 2021).

Unique proprietary data:

- Largest global dataset of HCM data

Kronos

Kronos primarily offers employee time tracking and scheduling services to businesses. The company processes data for more than 33,000 companies and 40 million employees, globally (Kronos, 2019).

Unique proprietary data:

- Mid-sized firms tend to be represented in the Kronos database (Chatty et al., 2020)

Homebase

Homebase processes hiring and onboarding, and time and attendance data for 100,000 plus small businesses.

Unique proprietary data:

- Average size of clientele is 8.4 employees (Chetty et al., 2020)
- 64 percent of available data are from restaurants and 15 percent are from retail stores (Chetty et al., 2020)

ICIMS

ICIMS is a leading applicant tracking system that offers a single recruiting platform that works across the hiring pipeline, from attracting talent, engaging and hiring potential new employees, to advancing those employees during their time at work. They are used in 10 percent of Global 2000 companies and 40 percent of Fortune 100 companies (ICIMS, 2021).

Unique proprietary data:

- Largest ecosystem of integrated recruiting technologies and services

Limitations of HCM data

Like job posting and work history data, HCM data do not reflect the full universe of a labor market as not all businesses use third parties for HCM. As mentioned above, HCM solution services only has a 71 percent global penetration (Statista, 2021). Certain industries and businesses like agriculture/farming and small family-run operations are likely to be underrepresented. Self-employed individuals, contractors, and even gig workers are also likely underrepresented in these data. Intermediaries' data cannot capture cash-based arrangements or manual check payouts – both of which remain prevalent. These differences may vary or be consistent across venders and other differences may be more pronounced among certain venders.

ONLINE LEARNING PLATFORM DATA

Technology is revolutionizing labor markets: changing the skill requirements of jobs, supplanting roles, and creating new occupations and hybridizing others. These technological changes are causing a schism between the skills workers have and the skills employers need. According to estimates:

- 54 percent of employees will need some form of retraining by 2022 (WEF, 2018a).
- 83 percent of companies have trouble finding skilled candidates (SHRM, 2019).
- 46 percent of organizations report an increasing skills gap (CompTIA, 2017).
- 79 percent of the world's CEOs are concerned that a lack of essential skills will threaten their organizations (PwC, 2019).

These concerns have led to the idea that workers will have to become lifelong learners to keep up with the expected ever-changing landscape of work (APEC, 2017; APEC, 2021). The concept of lifelong learning is leading traditional education and training providers, employers, workforce groups, and other stakeholders to reimagine job training. Particularly, how it can be facilitated efficiently and by which entities (online platforms, traditional facilitators, employers). To this end, online training platforms have gathered a wealth of data around how people acquire new skills.

Information extracted from online learning platforms

- Most and least effective teaching methods by subject matter.
- The skills people are learning, who is learning them and where.

The granular data online learning platforms collect can inform the types of incentives and feedback loops that encourage and support learning. The data can reveal what people are choosing to learn, how they perform and how outcomes can be improved. Competitive analyses can be performed around the competencies being acquired across regions. Temporal analyses can be performed to determine whether certain political, economic, professional or personal events affect when people are most likely to seek new skills and complete training. Table 3.5 and Table 3.6 provide an overview of coverage and sources of data from various online learning platforms.

Table 3.5 Coverage of online learning platform data

Third party intermediary	Coverage	
	How to access data	**Has global data?**
Coursera	By Request press@coursera.org	✓
EdX	Open-source https://edx.readthedocs.io/projects/devdata/en/latest/	✓
Udacity	By Request media@udacity.com	✓
Skillshare	By Request press@skillshare.com	✓

Table 3.6 Sources of online learning platform data

Third party intermediary	Source of Data				
	Competency training	**Knowledge sharing**	**User surveys**	**Communication**	**Collaboration**
Coursera	✓	✓	✓	✓	✓
EdX	✓	✓	✓	✓	✓
Udacity	✓	✓	✓	✓	✓
Skillshare	✓	✓	✓	✓	✓

Coursera

Coursera specializes in traditional academic-style offerings. It touts nearly 3,000 courses, and partnerships with over 200 universities and companies to develop them. The company reports having 82 million global learners including more than 6,000 campuses, businesses and governments.

EdX

EdX specializes in traditional academic-style offerings. It was founded by Harvard University and Massachusetts Institute of Technology and is known for its high-end lineage and partnerships. It reportedly has 25 million global learners and more than 3,000 courses.

Udacity

Udacity specializes in teaching digital skills and brands itself as an alternative to tech bootcamp-style learning. It reports more than 160,000 students across 190 economies.

Skillshare

Skillshare specializes in teaching creative skills and brands itself as offering practical skills that can be utilized right away. It reportedly has over 30,000 offerings.

Limitations of online learning platform data

Data gathered by online learning platforms will be limited to people and regions with internet access. It cannot inform learning styles, incentive structures or outcomes of traditional classroom learners. Datasets may be skewed toward learners from particular industries, professions and with certain levels of educational attainment.

Box 3.1 Online gig economy data

Intermediaries that operate within online labor markets or the gig economy create unique opportunities for researchers to analyze self-contained markets and conduct natural and field experiments. Platforms collect granular data about who takes gig jobs, where workers are located and the kinds of job workers are willing to accept and with what incentive. These platforms can provide insight into the factors that increase worker performance and reputation. These data can inform what political, economic, environmental, social, local market conditions, etc. compel workers to participate in the gig economy. It can also inform whether gig work creates new career pathways. Access information is provided in Table 3.7.

Uber is a rideshare platform that allows licensed drivers with relatively newer vehicles to operate as drivers for hire. Some 10 billion rides by 4 million drivers in 63 economies have been transacted through Uber, according to the company.

Instacart is an online delivery platform that allows drivers/workers to deliver groceries for hire. The service covers 85 percent of the US and 65 percent of Canada.

TaskRabbit is a platform that allows workers to offer their unique skills for hire (e.g., carpentry, plumbing, gardening, cleaning, moving/hauling, etc.). The platform is geared toward workers with vocational skills.

Upwork also allows workers to offer their unique skills for hire, but is geared toward professional services (e.g., software developing, video editing, writing, business consultation, etc.).

Limitations of online gig economy data

Data from online labor markets face shortcomings that may over- or underrepresent workers from certain socioeconomic backgrounds and regions. Data from certain vendors may be more robust in metropolitan areas compared to suburban and rural areas. For example, suburban and rural areas tend to have higher rates of vehicle ownership and may rely less on rideshare and delivery services. Workers in certain regions may also be more or less likely to engage in task-related gig work depending on the opportunities of the proximate labor markets.

ADDITIONAL SOURCES OF BIG DATA

Knowledge sharing and communication platforms can provide insight into how workers choose to collaborate and delegate tasks, how they work together to develop and troubleshoot products, how they learn new skills, how non-monetary rewards help garner contribution and what subjects have the most traction by region.

StackExchange hosts 173 knowledge share platforms, including Stack Overflow, the largest knowledge share platform for IT. The company reports 432 million monthly visits, 3.2 million questions asked, 3.5 million answers submitted and 13.5 million comments.

Quora is a knowledge sharing platform that operates like a social network. Visitors create an account and can follow topics and other users of interests. The site reportedly has 300 million monthly visitors.

GitHub is an open source repository that allows version-controlled editing of a number of file types, including code, Office documents, videos, etc. It claims to have more than 200 million repositories from some 65 million individual contributors and three organizations.

Slack is a tool designed for businesses that aims to bring communication and collaboration together in one place. It reportedly has 12 million users.

Investment tools that make research and data available to the public can provide competitive insight into how economies are investing their monetary resources. This data informs researchers about the industries and technologies that are garnering the most investment or that are becoming increasingly important within or across economies.

Capital IQ was developed by Standard and Poor's (S&P) and provides real-time investment and financial research metrics by collecting and analyzing some 135 billion data points.

Crunchbase is an open source investment research platform that allows users to contribute information about and/or research companies, investment funds, individuals, etc. around the world. The site allows user to track records, mergers and acquisitions, patents, fund raising campaigns, etc.

Table 3.7 and Table 3.8 provide a brief overview of coverage and sources of data from the gig economy and additional platforms described earlier.

Table 3.7 Coverage of gig economy and additional data sources

Third party intermediary	Coverage	
	How to access data	Has global data?
Uber	By Request press@uber.com	✓
Instacart	By Request press@instacart.com	US & Canada
TaskRabbit	By Request press@taskrabbit.com	US, UK & Canada
Upwork	By Request press@upwork.com	✓
StackExchange	Open source data.stackexchange.com/	✓
Quora	By Request help.quora.com/hc/en-us/requests/new	✓

GitHub	By Request press@github.com	✓
Slack	By Request slack.com/help/requests/new	✓
Capital IQ	Paid https://www.spglobal.com/marketintelligence/en/solutions/sp-capital-iq-platform	✓
Crunchbase	Open source data.crunchbase.com/docs/daily-csv-export	✓

Table 3.8 Sources of gig economy and other data

Third party intermediary	**Sources**				
	Gig economy	**Knowledge sharing**	**Communication and collaboration**	**Investment**	**User survey**
Uber	✓	×	×	×	✓
Instacart	✓	×	×	×	✓
TaskRabbit	✓	×	×	×	✓
Upwork	✓	×	×	×	✓
StackExchange	×	✓	✓	×	✓
Quora	×	✓	✓	×	✓
GitHub	×	✓	✓	×	✓
Slack	×	×	✓	×	×
Capital IQ	×	×	×	✓	×
Crunchbase	×	×	×	✓	×

STRENGTHS OF ADDITIONAL BIG DATA SOURCES

As discussed earlier, alternative big data sources include:

- Online job ads and employee work history data
- HCM data
- Online learning platform data
- Online gig economy data
- Knowledge sharing and communication platforms

The near ubiquitous digitalization of the HCM process, from online job postings to applicant–employer matching, hiring and onboarding, time tracking, payroll management and other processes, has created scores of rich alternative sources of labor market data. Research has shown how combining these new data sources with traditional sources can augment our understanding of labor markets, while solving for some of the barriers associated with traditional sources.[2]

Publicly funded labor market data suffer from survey bias and non-response bias, are produced quite infrequently, capture only a moment in time, are costly and lengthy to produce and do not allow for granular insights. In contrast, alternative sources capture the natural behavior of employers, job seekers and workers (Faberman and Kudlyak, 2016; Turrell et al., 2019); the data are collected daily and processed in real time (Carnevale et al., 2014; Horton and Tambe, 2015; Faberman and Kudlyak, 2016; Cajner et al., 2019; Turrell et al., 2019; Bartik et al., 2020;

[2] See Carnevale et al. (2014), Horton and Tambe (2015), Faberman and Kudlyak (2016), Cajner et al. (2019), Turrell et al. (2019), Bartik et al. (2020), Cajner et al. (2020) and Zhang et al. (2021).

Cajner et al., 2020; Zhang et al., 2021); have sample sizes that at least mirror traditional sources at costs that are comparatively negligible (Horton and Tambe, 2015; Cajner et al. 2019); and allow for granular analyses (Carnevale et al., 2014; Horton and Tambe, 2015; Turrell et al., 2019). In Table 3.9, the main differences in traditional labor market data and big data are outlined.

Table 3.9 Comparison of traditional labor market data and big data

Type of data	Years of data	Ease of time series analyses	Data representative - ness	Compatibility across economies	Real-time data access	Regular taxonomy updates	Data granularity
Traditional labor market data	~50	High	Apply statistical sampling methods and weights	✓	×	×	Low
Big data	~10	Medium	Captures digitized labor market; can benchmark against public data to gain insight	✓	✓	✓	High

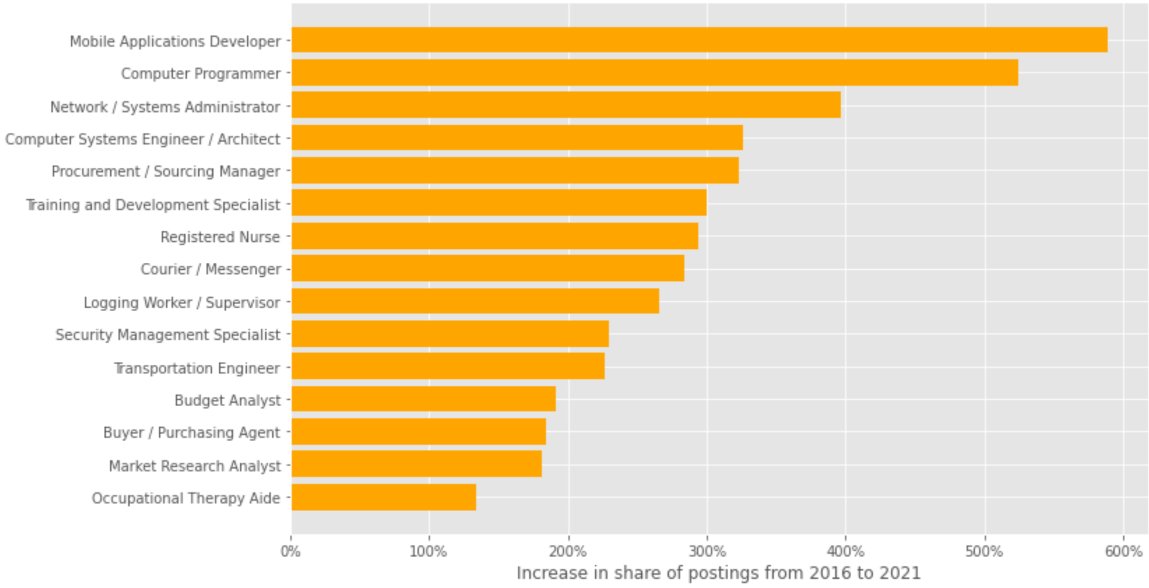
Level of granularity: Traditional job vacancy and employment surveys are not well-suited for granular analyses. Sample size may quickly become an issue when analyzing crosscuts or subsets of microdata (Horton and Tambe, 2015; Faberman and Kudlyak, 2016; Turrell et al., 2019). Economies tend to keep their taxonomies broad and rigid to mitigate sampling issues and for longitudinal consistency. This hinders economy-wide and regional insights into emerging, fast-growing and declining job titles; fails to elucidate the skills and tasks required for specific jobs (as opposed to occupations as a whole); and fails to capture rapidly changing skill demand. Disparate practices around survey anonymity impede micro-analyses of employers and under- or over-represented groups (e.g., race, ethnicity, gender, etc.).

Intermediaries such as Indeed, LinkedIn, EBG and Monster aggregate data on workers from résumés uploaded to the internet, social profiles created on job boards, and online job ads posted by employers. These data capture daily changes in the supply, skillset and educational attainment of talent; and occupational demand, skill demand, wage offerings, and educational and experience requirements across industries and employers (Carnevale et al., 2014; Horton and Tambe, 2015; Faberman and Kudlyak, 2016; Turrell et al., 2019). This allows for real-time macro- and micro-level analysis of labor market trends and shocks across and within economies, often down to the city level. An example of this kind of analysis involves looking at the top emerging occupations across an economy. Figure 3.1 compares percent change in occupation posting share from 2016 to 2021 to determine the occupations that have grown the fastest.

Similarly, economies can gain insight into top-growing skills as well. Figure 3.2 compares skills demand in 2016 and 2021 to determine macro emerging skill trends

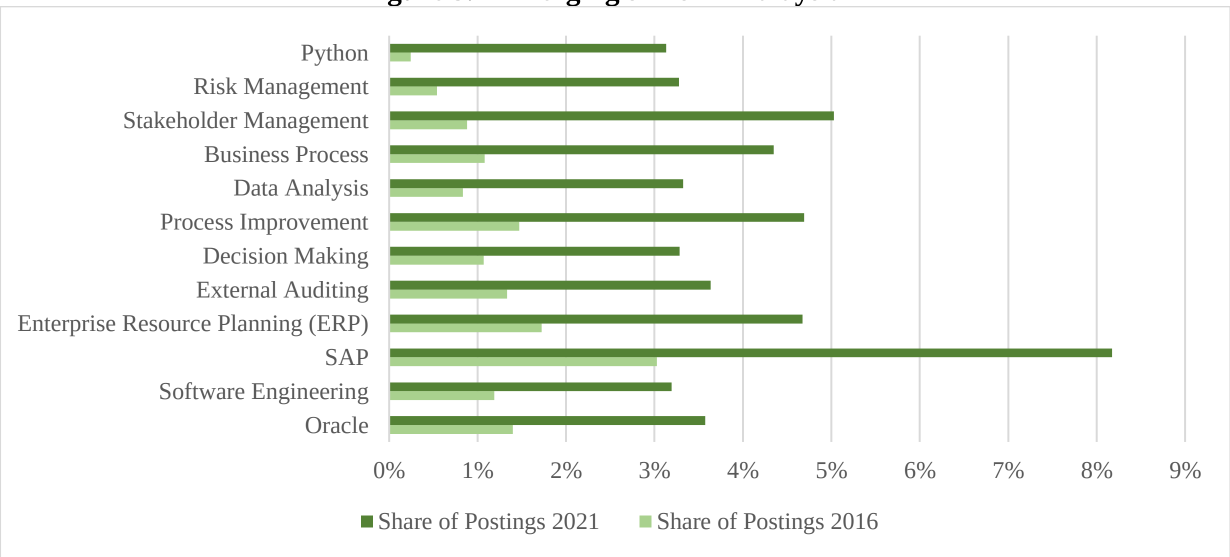
HCM solutions providers like ADP, Kronos, and Homebase capture labor market microdata as a by-product of their primary business operations. These HCM providers process payroll, benefits, time and attendance, talent management, and other HCM and business operations data. These data have served as proxies for analyzing layoffs, hiring, and labor shocks in real time (Cajner et al., 2019; Bartik et al., 2020; Cajner et al., 2020). The data are also used to analyze pay and employment inequities, turnover and retention, and labor market shocks, among other subjects.

Figure 3.1 Fastest growing occupations in Malaysia



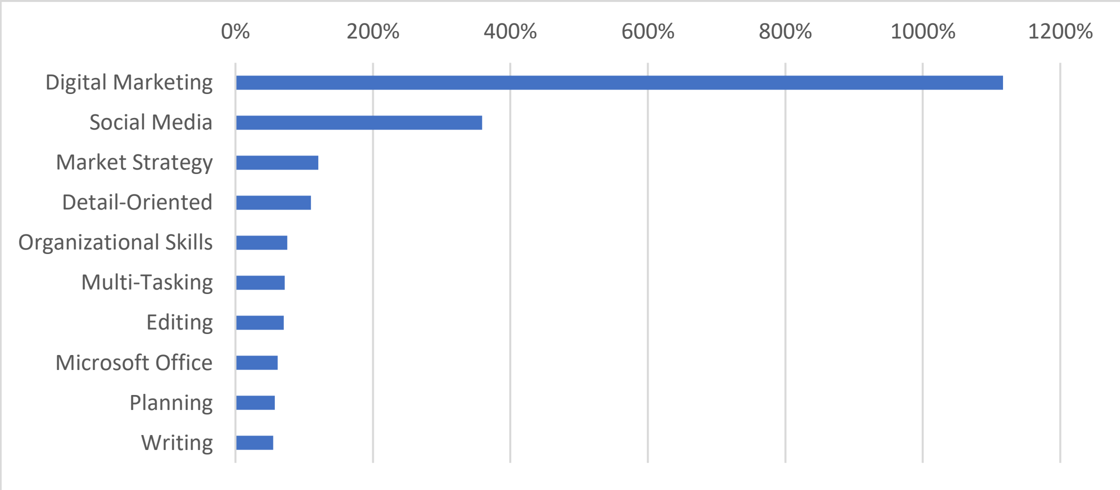
Source: Emsi Burning Glass (EBG) analysis.

Figure 3.2 Emerging skills in Malaysia



Source: EBG analysis.

Figure 3.3 Top skill changes for marketing specialist, 2010–2020



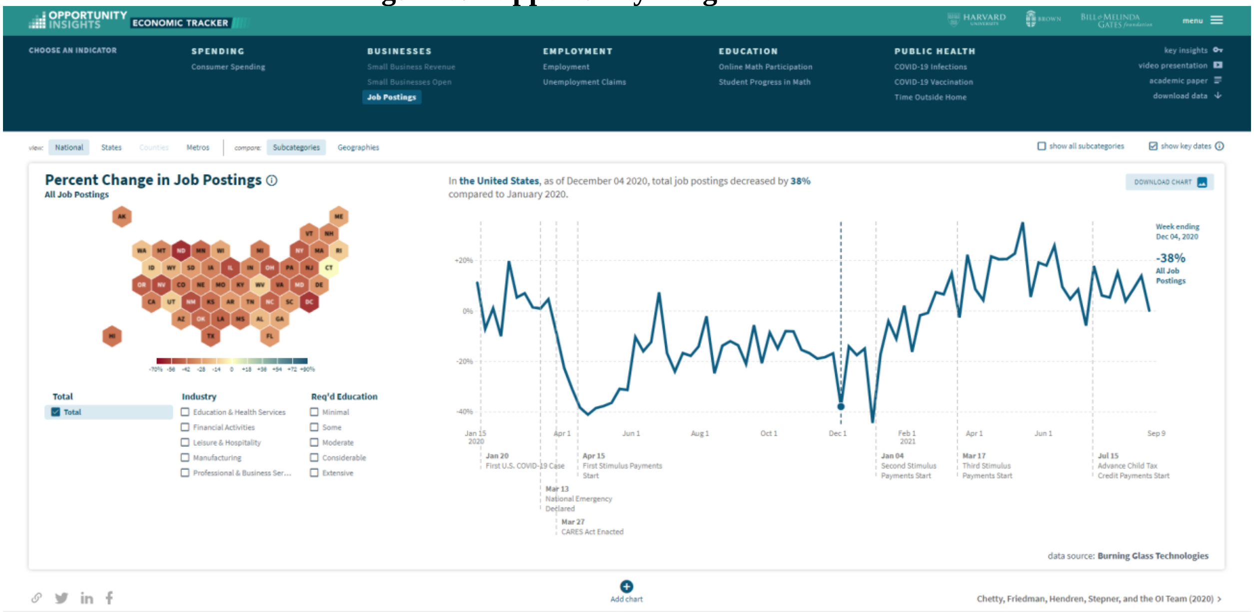
Source: EBG analysis.

One example of the benefits of increased granularity is being able to look at top skills associated with an occupation over time. Figure 3.3 shows the fastest growing skills for marketing specialists based on job posting data from 2010–2020. Digital marketing and social media have grown the fastest in this time period and have radically changed the way marketing specialists perform their job functions. Without understanding the skill changes within a job over time, this type of analysis would be impossible. This analysis can help policymakers working on reskilling understand the skills they should focus on and those that they should stop training on, for example.

Quick to collect and update data content and account for shocks: These alternative data sources serve as precursors or early indicators for lagged traditional sources. Many of these data sources update daily or weekly rather than monthly, quarterly or annually as is the case with most traditional sources. This cadence means that microtrends or shocks can be seen in these sources almost immediately. This timeliness is vital for studying issues such as the COVID-19 pandemic, which wreaked havoc on economies quickly. In combination with unemployment data, tracking real-time job posting data helped researchers understand what was going on in the economy at a detailed level immediately (Chetty et al., 2020; Forsythe et al., 2020). An example of this is in Figure 3.4 which shows the Opportunity Insights tracker. This tool was developed by an organization at Harvard University to provide a real-time picture of the US economy amid the COVID-19 pandemic.

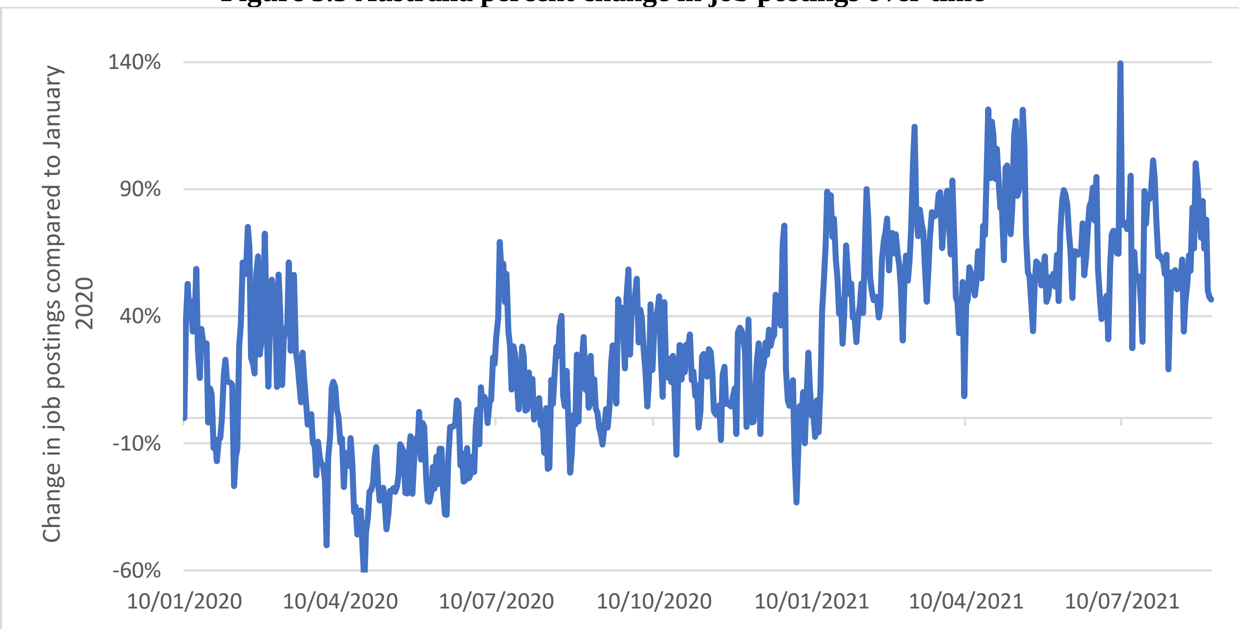
Policymakers can use a similar view to the Opportunity Insights tracker to see real-time how job postings change during COVID-19. Figure 3.5 and Figure 3.6 show examples of job postings over time from other economies (respectively, Australia and New Zealand) from January 2020 to August 2021.

Figure 3.4 Opportunity Insights tracker



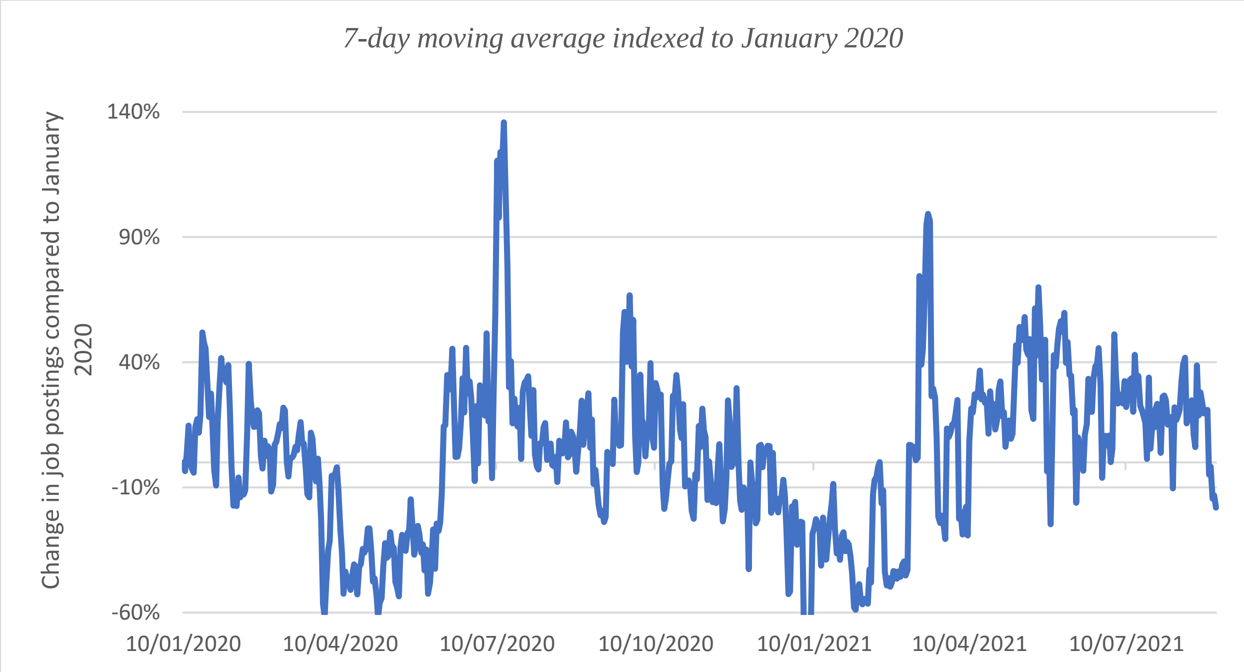
Source: Opportunity Insights (n.d.).

Figure 3.5 Australia percent change in job postings over time



Source: EBG analysis.

Figure 3.6 New Zealand percent change in job postings over time



Source: EBG analysis.

The data in Figure 3.5 and Figure 3.6 show that, in Australia, job postings declined in the 2nd quarter of 2020 but have been on an upward trend since then. In contrast, in New Zealand, there were two dips of note in job postings: in the 2nd quarter of 2020 and the 1st quarter of 2021. These findings could be matched with other policy or events data – such as spikes in COVID-19 daily reports cases or the implementation of restrictions on movement – to determine the causal impacts of these events on jobs.

Researchers have found that, relative to traditional labor market data, real-time data more accurately capture the breadth and depth of shocks. Bartik et al. (2020) found that the Current Employment Statistics (CES) and Current Population Survey (CPS) produced by the US failed to adequately capture the magnitude of the shock because of the survey collection schedule. Using CES and CPS data, they report, from late January to early July, overall employment appeared to have only declined 18 percent. However, using Homebase and Kronos data, they found a 60 percent decline in employment among small businesses and a 35 percent decline among large businesses, respectively. These alternative sources provided an early look at the disparate impact on small businesses compared to large businesses, that government data was unable to capture right away.

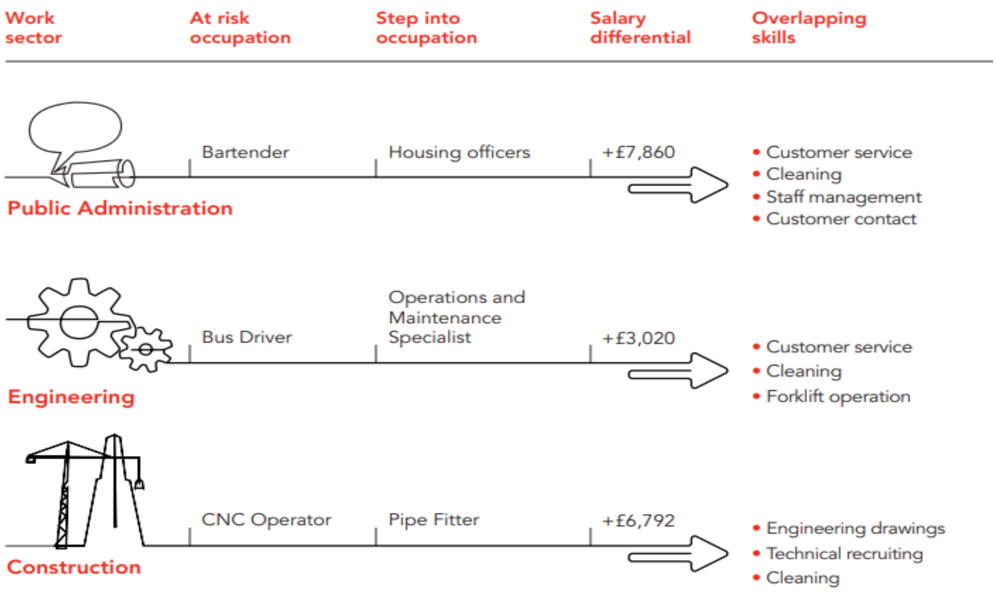
In a similar macro-analysis of the impact of COVID-19 pandemic on the US labor market, Cajner et al. (2020) operationalized real-time payroll data from ADP to find the change in hiring and paycheck distribution amid the crisis. They found some 13 million workers lost their jobs in just two weeks from 15 March to 28 March 2020, whereas data from unemployment claims captured about 3 million less unemployed workers. City & Guilds and Burning Glass Technologies (2021) leveraged EBG real-time microeconomic data to retool workers who were hardest hit by the pandemic in the UK. The data identified hard-hit occupations, the skills required to perform those jobs and mapped skill-based career pathways (that required little to no training) to jobs that were stable or growing in the midst of the pandemic, while still offering a livable wage (see Figure 3.7 and Figure 3.8).

Figure 3.7 At-risk occupations

Top 10 declining at risk occupations

At risk Occupation	Growth from Sep19/Jan19 to Sep20/Jan20
Pet Care Manager	-93%
Bus Driver	-83%
Nanny / Babysitter	-72%
Veterinary Nurses and Assistants	-72%
Busser / Banquet Worker / Cafeteria Attendant	-71%
Fundraising / Development Specialist	-61%
Bartender	-60%
Survey Researcher	-58%
Executive Assistant	-55%
Telemarketer	-53%

Figure 3.8 Career pathways for at-risk occupations



For Figure 3.7 and 3.8
Source: City & Guilds Group and Burning Glass Technologies (2021).

The magnitude of data captured also means that, despite volatility, there is generally a substantial sample size to use for smaller time periods. Billions of observations are captured and processed daily by these intermediaries and transformed into robust databases. The following is a glimpse of the magnitude of the real-time data being parsed:

- 60 to 70 percent of US job openings are posted online (Carnevale et al., 2014)
- 76 percent of unemployed and 33 percent of employed US workers were seeking new employment online by 2011 (Faberman and Kudlyak, 2016)
- Over 400,000 unique job postings are parsed by EBG in Australia and New Zealand annually

Quick to update taxonomies: Traditional sources of labor market data are slow to update taxonomies, which impacts the ability to understand the current (and future) labor market. One salient example of this is with the occupation 'data scientist'. As early as 2012, publications like the *Harvard Business Review* were referring to the position as 'the sexiest job of the 21st Century' (HBR, 2012). However, data scientists were still lumped in with 'mathematical technicians' in the US Bureau of Labor Statistics data until a new classification split them into their own occupation in 2018. Much of the publicly available data from the Bureau of Labor Statistics still do not reflect the new taxonomy (as of July 2021).

In Canada, an occupation must meet certain requirements, such as a minimum of about 5,000 workers and a list of duties that show that the occupation is significantly different than other existing occupations. In addition, new occupations require a list of job titles and employment requirements used by a majority of employers to define it. Often, emerging occupations do not meet those requirements. That said, data scientists are represented by the National Occupation Classification (NOC) unit group 21211 in NOC 2021.[3]

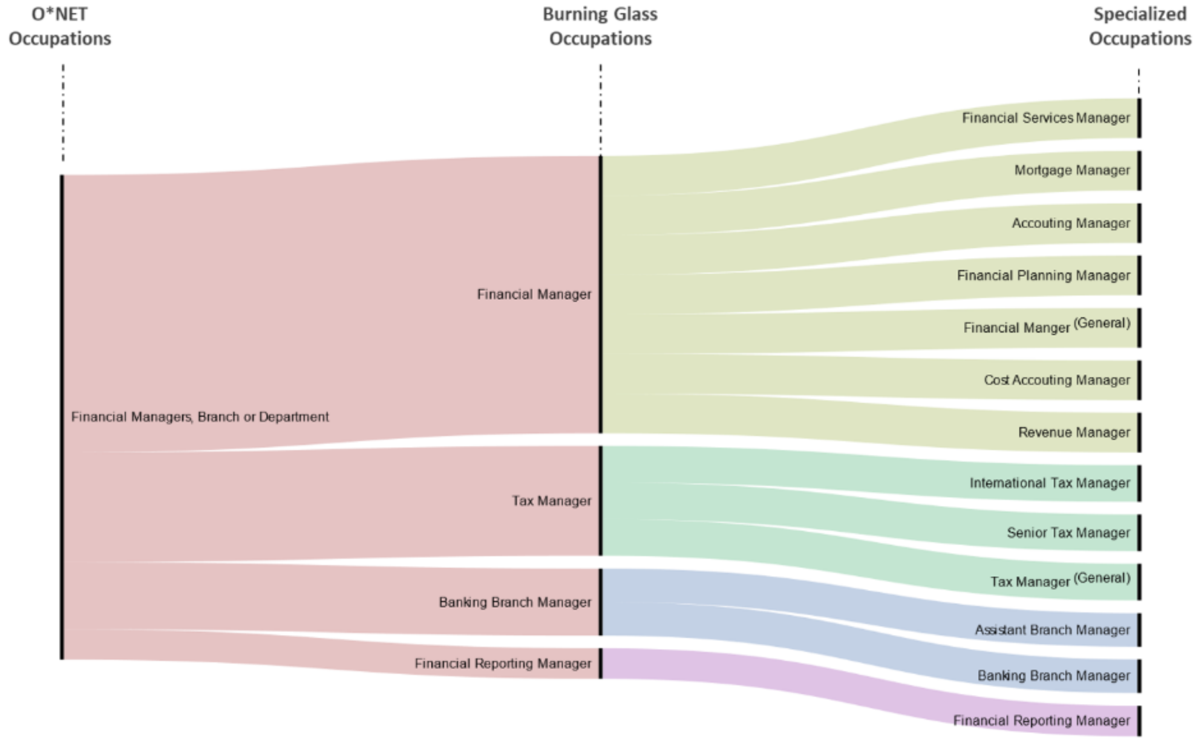
Other APEC economies, meanwhile, still have not adapted their occupation classifications to separate data scientists (as of 2021). Similar stories are true of other emerging jobs, such as UI/UX designer, e-commerce analyst, social media strategist, and others.

In contrast, the granularity and speed at which big data providers collect and process their data provide a means for rapid response and tracking. They can track changes in job titles and clusters of skills and more accurately characterize the labor market. Figure 3.9 shows how standard O*NET occupations for financial managers are split into more specific occupations that allows for a much deeper understanding of labor market demand and supply. EBG started collecting data on data scientists as early as 2014. This meant that policymakers could understand quickly the ways in which skills were being combined, and they could focus education policy, training provision and other workforce development on the right combination of skills.

This level of granularity allows for a better understanding of available career pathways, as skills in some occupations might differ more than expected. For example, the top skills for mortgage manager include mortgage lending, underwriting and people management, while the top skills for a cost accounting manager include general accounting, financial reporting and financial analysis. Breaking these out into separate occupations means that career pathways and salary bands can be based on these job-specific skills rather than generic financial manager skills, which would obfuscate differences between these occupations.

[3] See Canada's 2021 National Occupational Classification: https://noc.esdc.gc.ca/Structure/Noc2021.

Figure 3.9 Sample of big data occupation taxonomy



Source: EBG document.

HOW BIG LABOR MARKET DATA HAS BEEN USED

Decision makers in academia, workforce development, private companies, and other organizations have leveraged third-party labor market data for years to make informed decisions and policies. These data have been utilized in a number of ways, including to: shape training programs and curricula, identify future of work trends like growing and waning occupations and skills, future-proof labor markets and corporate workforces against disruptive occupations and skills, model career pathways to lead workers to better opportunities, assess skills gaps, and augment worker–employee matching.

Skill projection analysis

In 2017, IBM and the Business–Higher Education Forum used data from online job advertisements to conduct a skill projection analysis to understand how data science and analytics were reshaping the US labor market (Miller and Hughes, 2017). The analysis posited that traditional labor market data created an ‘information gap’ due to a taxonomical system that is updated infrequently (every 10 years in the US). This hampers the ability to identify emerging trends as they occur. The pair mined 130 million online job ads from EBG for their period of analysis and analyzed how 300 analytical skills were reshaping the US labor market.

The data allowed Miller and Hughes (2017) to conduct longitudinal analyses to estimate the five-year growth in demand for jobs requesting data science and analytics skills (15 percent growth from 2.4 to 2.7 million jobs). The data showed a talent shortage had already existed as the average time to fill these roles were higher than market average, and salaries were inflated relative to all bachelor's and graduate-level jobs. On a more granular level, they were able to identify the industries that were embracing these skills and how the skill requirements varied across industries. They also noted that following industry penetration, there was increased skill, education and experience requirements across several affected occupations. Miller and Hughes (2017) implore education and job training providers to account for the increasing demand for analytical skills to avert an exacerbated shortage.

Intermediaries that track online job ads like EBG conduct custom skill projection analyses to forecast the growth of skills across regions and industries up to 10 years out. The data allow these analyses to be as granular as tracking individual skills or specific clusters of skills. For example, Salesforce contracted EBG to identify the historical, current and projected growth in demand for knowledge of its software (Vilvovsky et al., 2020). Similarly, video game producer, Epic Games, launched a study to identify the value and projected demand for 3D graphic skills across 25 labor markets (Vilvovsky et al., 2021). Markow and Sederberg (2020) leveraged the data to detect and project the future growth of the most disruptive emerging skills and technologies.

Labor market disruptions

Like in the above scenario, the taxonomical limitations of traditional labor market data make alternative sources astute options for identifying technological disruptions. Acemoglu et al. (2020) and Jiang et al. (2021) also operationalize EBG data to measure the impact of artificial intelligence (AI) and financial technology (fintech) on the US labor market, respectively. They assess the impact on industries, occupations, workers and the broader labor market. Both studies conclude industries and companies with high exposure to AI and fintech saw a reduction in hiring, evidenced by a decrease in online job posts.

The granularity at which job postings data are available allowed Acemoglu et al. (2020) to exclude AI-*producing* firms from their analysis and only focus on AI-*using* firms. The study found increasing demand for AI skills as early as 2010, with a significant acceleration in demand around 2015–2016. They also found early signs that AI was changing work. The data showed increased exposure to AI changed the skills and tasks listed in job posts. Acemoglu et al. (2020) suggest this as evidence that automatable tasks were no longer being performed or required. However, they found new tasks and skills being requested. The study determined that AI-infused firms and industries reduced hiring of non-AI roles, but the trend had no effect on the wider labor market.

Similarly, the granularity of job posting data allows Jiang et al. (2021) to pinpoint the kind of worker that is most negatively impacted by the rise of fintech. After analyzing some 162 million job posts, they determined occupations with average salaries and an intermediate education requirement (high school to some college) were most negatively impacted by fintech. This is compared to occupations with no education requirement or that require a bachelor's degree or higher. The study found that occupations with the most exposure saw a 5 percent decrease in job postings over 11 years. They were also able to pinpoint the most disruptive sets of skills – fintech, data analysis, blockchain, and robo-advising. The study found that a rise in fintech is associated with a decrease in employment growth and an overall increase in education, experience and complex skill requirements (especially software skills) across exposed firms.

However, these effects were concentrated within the financial industry and limited to large financial hubs like New York City, Seattle, Chicago, etc.

Labor market disruption can also be analyzed through the lens of labor supply data. LinkedIn Economic Graph used LinkedIn member profiles to analyze AI talent trends in the European Union (LinkedIn, 2019). The goal of this analysis was to investigate geographic, industry, demographic and mobility trends. The methodology used in this analysis was made possible by the level of granularity of the data. First, AI talent was identified by searching individual LinkedIn profiles for a set of AI keywords. Then, a model was applied to refine this group such that the sample had the necessary experience and skills. The researchers found that AI talent was not distributed the same across the European Union: the UK, France and Germany supplied half of the AI talent in the region. They also found that AI talent was mostly comprised of men, and women made up only 16 percent of the AI talent pool. The granularity of the data also allows economy-level comparison by industry, employer type, degree type and career length.

These reports are just a sample of how alternative labor market data can provide foresight to policymakers, workforce groups, and academia – allowing ample time for preemptive responses.

Curriculum development

The Fourth Industrial Revolution is marked by rapid technological advances and ever-emerging new technologies. This is ushering in a period of lifelong learning – putting pressure on workers, training providers and labor markets to keep up. Critics point to widening skills gaps as evidence that traditional academic institutions and workforce training groups are unable to retool quickly enough to supply talent. This has given rise to a wave of new online learning platforms, and even employer-provided training, that aim to fill these gaps. However, the lack of a common framework for teaching emerging competencies has created asymmetric information around what is actually being taught, and is disadvantaging labor markets. Talent does not have clear signals about the specific competencies and skills they should learn. Employers are uncertain if micro-credentials and other competencies listed on résumés match their needs. Training providers not only lack insight into the skills employers are requesting but are also unsure if the knowledge prospective students have is in line with program requisites. Gromov et al. (2020) and Kitto et al. (2020) demonstrate how an off-the-shelf software from a third-party labor market intermediary might solve these asymmetries.

EBG developed a course content tagger that leverages natural language processing (NLP) to ingest curricula and output the skills taught in courses. The content tagger utilizes EBG’s skills database of over 17,000 skills to automate the comparison of course descriptions. The company’s skills ontology was developed, and is constantly updated, based on the skills and competencies listed in job postings. Gromov et al. (2020) demonstrate how this process can inform curriculum development. The researchers used the content tagger to extract the skills taught in select analytics courses at the University of Technology Sydney. They then used EBG’s data to identify the top skills for select data science and analytic occupations in Australia. They found the courses covered less than 50 percent of the skills requested in the market. Kitto et al. (2020) demonstrate how the content tagger can automate the comparison of courses and close the information gap around the competencies taught by different providers.

Job search and career pathways

Information asymmetries also stymies the job search process for job seekers and contributes to underemployment and employer–employee mismatch. Job seekers have limited information about how their skills translate to different occupations and career areas. This causes some seekers to narrow their job search thus limiting their options, while others might search too broadly. Additionally, unemployment insurance imposes temporal sanctions that incentivizes job seekers to accept jobs they would not otherwise take.

Belot et al. (2018) built an online job advice tool and used it to conduct a survey on 300 unemployed job seekers in the UK. The tool provided participants with a list of jobs they could reasonably perform based on their previous occupations and skills. This insight was derived from data on previously realized transitions and 'information on transferable skills across occupations' (Belot et al., 2018). The tool also provided a list of online job openings that represented about 80 percent of the vacancies in the UK. The researchers reported a statistically significant change in job seeker behavior. Narrow job seekers broadened their search, and seekers that started broadly began to focus their job search.

Career pathway analyses has the potential to support seekers with their job search: LinkedIn's Economic Graph team created a tool that uses LinkedIn data to map out career paths by comparing job seekers' skills to job titles called the LinkedIn Career Explorer. The tool allows users to find potential job transitions by taking the user's current role and finding skills in common with other roles using a skill similarity score. Career pathway analyses also has another important use: helping shape public and internal workforce policies. Australia (2019) used online job postings data to assess the skills Australians need to remain competitive amid the Fourth Industrial Revolution. It then utilized EBG's proprietary skill adjacency analysis to create a plan to transition workers facing the highest risk of automation to jobs with lower risk. Skill adjacency analyses broaden the number of career pathways by leveraging job postings data and machine learning algorithms to identify occupations that are similar based on skills, competencies, credential, experience and salary. The World Economic Forum (WEF) explored two additional use cases. The analysis can save corporations time and money, and increase productivity, by providing a blueprint of the occupations that are disrupting a particular industry and the talent within a company that can quickly be upskilled to remain competitive (WEF, 2019). WEF (2018) identified roles traditionally filled by women and showed that a skill adjacency analysis reveals career pathways to male-dominated roles.

Measuring the efficacy of policy

Alternative labor market data can help measure the effects of fiscal and labor market policy on a granular scale. Cho (2018) leveraged ADP payroll data to glean insight on how the American Recovery and Reinvestment Act of 2009 (ARRA) impacted hiring, wages and hours worked across specific companies. The ARRA was a USD 831 billion stimulus package meant to pull the US economy out of the Great Recession. The granularity of ADP payroll data allowed Cho (2018) to measure how companies that received government contracts as part of the ARRA managed the demand shock. He found the rate of hiring was higher at companies that received contracts compared to companies that did not; this effect lasted two years. He also found wages increased at contract receiving firms.

Yildirmaz (2016) further demonstrates the granular insights that can be drawn from HCM data. Also leveraging ADP payroll data for the US, Yildirmaz (2016) seeks to understand wage growth and turnover as ADP's payroll data tracks workers across companies. The report finds workers who switch jobs see a 6 percent wage increase compared to a 4.6 percent increase among workers who remain at the same job. Yildirmaz (2016) is also able to identify the industries, regions and demographics of workers experiencing the most wage growth and turnover. As economies around the world seek to address pay gaps and other labor market disparities, these types of granular data about workers can help measure the efficacy of such policies.

4. RECOMMENDATIONS FOR INTEGRATING NEW DATA SOURCES AND COLLABORATION METHODS

This section reviews the options available to APEC economies – including APEC as an organization, member economy governments, academics and other stakeholders – to integrate new data sources into their labor market information systems and processes. We will provide recommendations for overcoming challenges related to the data, the modes of collaboration and potential use cases.

ADOPTION OF ALTERNATIVE DATA SOURCES AND OVERCOMING DATA CHALLENGES

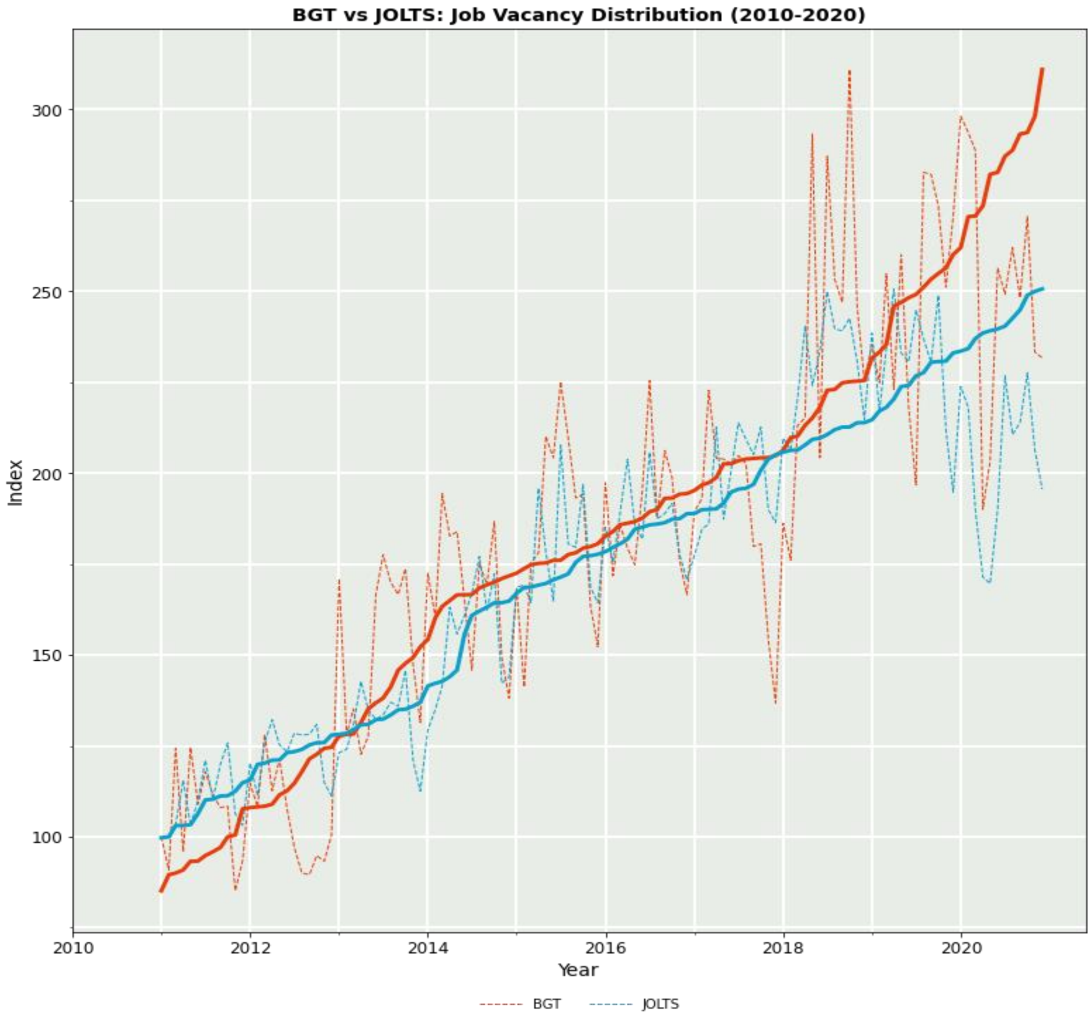
In this section, we explore some of the main challenges stakeholders face in adopting big data sources to be used in analyses, and provide methodological recommendations to overcome these hurdles and get the most out of big labor market data.

Length and stability over time. Due to the nature of digital data sources and relatively recent advancements in technology, many big data sources have been collecting data for less time than their traditional counterparts. This means that broad shifts in the nature of the labor market from more than a decade ago are difficult to capture with newer sources of data. However, these alternate data sources offer a richer and more real-time picture of current happenings in the labor market.

Big data sources are often less stable over time in terms of use for time series analyses. As improvements are made in scraping and processing technologies, the universe of data that can be accessed by intermediaries changes and grows, improving over time. The result is that the data becomes more representative of the general population, but also shifts in terms of what is captured. In order to perform time series analyses, looking at the distribution of occupations, skills or industries rather than the absolute number of data points can help control for some of these changes. This type of data is more useful than traditional data when trying to understand rapidly changing technologies, such as Python, Tiktok, cryptocurrencies, etc.

In the long term, trend analyses show alternative data have trends that mirror traditional data (see Figure 4.1). A comparison of job posting data to JOLTS data, a public survey dataset that asks employers about vacancies and separations, shows a correlation over time of 0.75-0.98 (Carnevale et al., 2014; Faberman and Kudlyak, 2016; Turrell et al., 2019; Cajner et al., 2019; Bartik et al., 2020; Cajner et al., 2020).

Figure 4.1 Comparing job posting data to JOLTS data

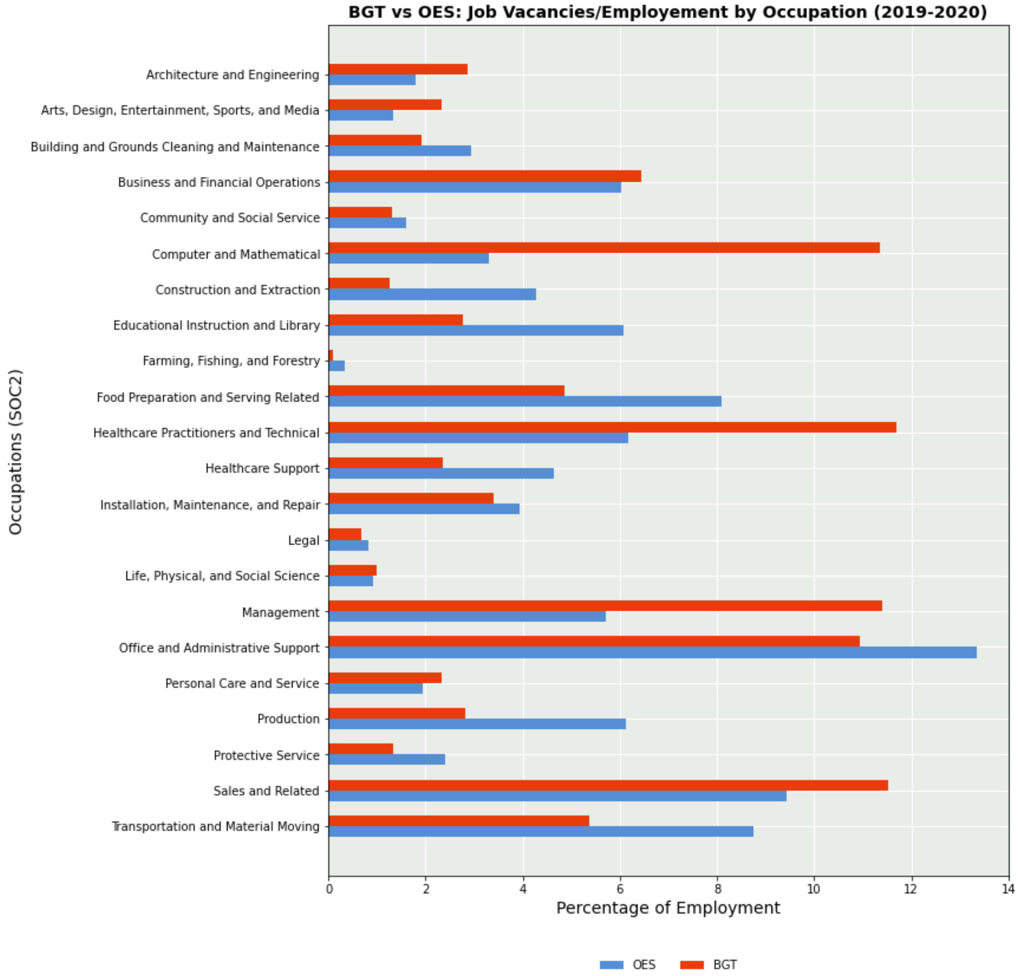


Notes: BGT refers to EBG job postings data. JOLTS refers to the Job Openings and Labor Turnover Survey conducted by the US Bureau of Labor Statistics. Each dataset is indexed to itself to allow for a comparison of trends rather than levels.
Source: EBG analysis.

Representativeness. The changes in the sources scraped over time also affect the representativeness of big data, as do limitations based on the digital landscape. Intermediaries that rely on scraped data sources may not capture parts of the labor market that are unlikely to be digitized. Those that rely on client or customer data are also likely to only cover certain sections of the labor market. Benchmarking the data against public data sources that are representative of the general population can provide stakeholders with an understanding of where these data can provide insight and where they fall short (see Figure 4.2).

For some analyses, especially those around digitalization and emerging technology, the bias of many of these large datasets can act to stakeholders' advantage. The data that are available in these sources are likely to be the most digital or advanced jobs in the market as a whole, which means that any trends around digitalization or advancement of digital skills found through big data are likely to be a conservative measure of the true digitalization (O'Kane et al., 2020).

Figure 4.2 Benchmarking big data against public data sources



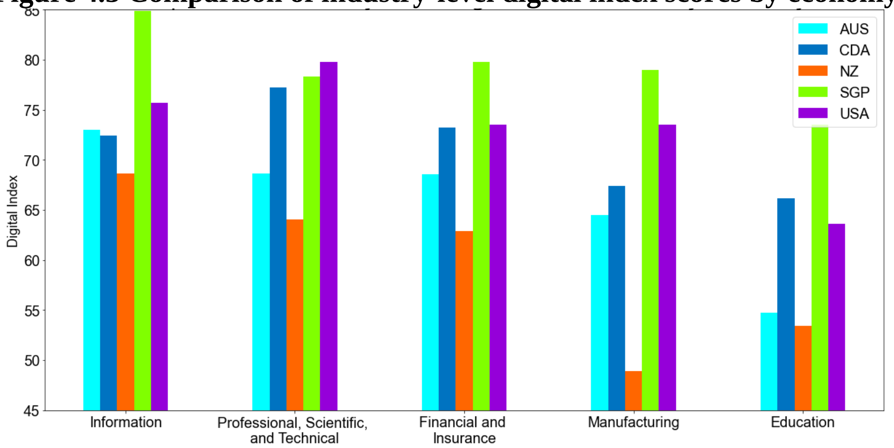
Notes: OES refers to Occupational and Employment Statistics as created by the US Bureau of Labor Statistics. BGT refers to EBG data.
Source: EBG analysis.

While big data may not be as representative of the general labor market, it does a better job of representing certain corners of the labor market, particularly those that are digitally intensive and rely heavily on technology. In order to gain a better overall picture, big labor market data can be used to complement traditional sources of data in understanding trends in these digital fields and in the labor market more broadly.

Compatibility. One of the primary benefits of working with many traditional data sources is the ability to compare data across economies due to standardization of taxonomies such as ISCO and ISIC at the occupation and industry levels and ESCO or O*NET at the skills level. This benefit is extended to many big data sources, which have specifically created their taxonomies to either mimic or be compatible with these standard taxonomies. However, in some cases, big labor market data variables are created using different, proprietary taxonomies that may not be directly linkable to other more traditional sources of data. Some intermediaries offer services to align taxonomies used by stakeholders, including higher education institutions, employers and governments, to their proprietary taxonomies to allow for further analysis. Stakeholders should consider the levels of compatibility required and the pros and cons of each taxonomy to understand the best possible use of these data.

Compatibility can also be easy to address when using big data sources and comparing across economies. APEC (2020a) employs both sources to measure and compare the digital skills gap across APEC economies. The report reveals, economy by economy and sector by sector, the degree and rate of employer demand for digital skills and the complexity (i.e., low complexity to high complexity digital skills). Figure 4.3 shows these comparisons for the Australia; Canada; New Zealand; Singapore; and the United States.

Figure 4.3 Comparison of industry-level digital index scores by economy



AUS=Australia; CDA=Canada; NZ=New Zealand; SGP=Singapore; USA=United States
Notes: The digital index is a score that allows for cross-economy comparison of the level of digitalization in jobs based on the skills required in job postings. Higher scores indicate higher levels of digitalization.
Source: APEC (2020a).

The data are also used to uncover the supply of workers with the necessary digital skills relative to job demand across each economy. Zhang et al. (2021) take LinkedIn and EBG data one step further to compare and contrast worldwide supply and demand for artificial intelligence (AI) skills in résumés, social profiles and job postings. Data from these intermediaries can be leveraged to identify where opportunities exist to move workers and close skills gaps.

Accessibility. While some big data sources are open source and free to use, others require purchase agreements or are limited in terms of availability. Moreover, most big data sources are just that – 'big'. Many data systems are not equipped to handle big data that updates with real-time frequency. The next section discusses ways to address these considerations and understand available methods of collaboration.

SELECT USE CASES BY POLICYMAKERS

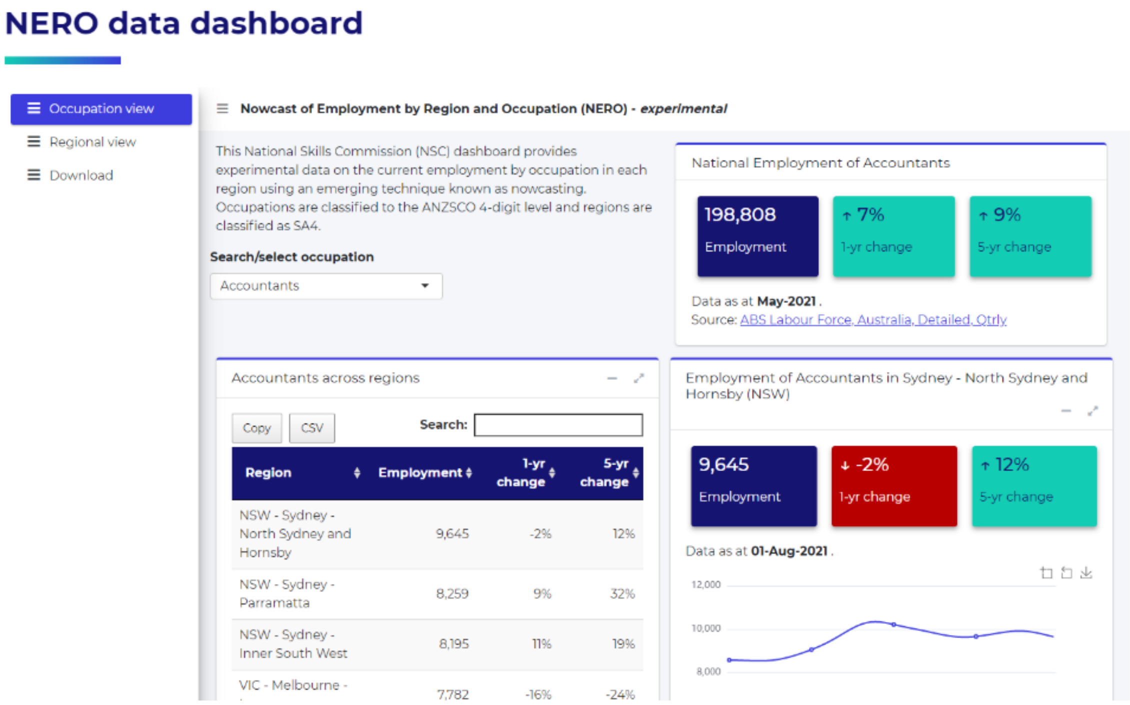
Utilizing real-time insights for quick policymaking

There are various examples of policymakers using big labor market data to increase granularity, timeliness and relevance of labor market information. The Australian government uses big data, including near real-time information from online job advertisements, and has worked with labor-market analytics firm EBG in many capacities. For example, the National Skills Commission of Australia uses big labor market data in their Jobs and Education Data Infrastructure (JEDI) tool. JEDI provides real-time data on the Australian labor market, helping policymakers to navigate the changing economy by providing data on the labor market, workforce changes, and current and emerging skills needs.

This kind of emerging skill analysis is not possible with traditional labor market data due to data lags and taxonomical changes occurring infrequently. The National Skills Commission also developed Nowcast of Employment by Region and Occupation (NERO) (see Figure 4.4) which directly addresses the issue of time lag in public data. Employment data in 355 occupations across 88 regions is only available every five years from the Australian Bureau of Statistics Census of Population and Housing, whereas this tool enables the user to understand monthly employment data. The tool implements an emerging technology called ‘nowcasting’, a method using traditional and real-time data to estimate timely and granular employment trends.

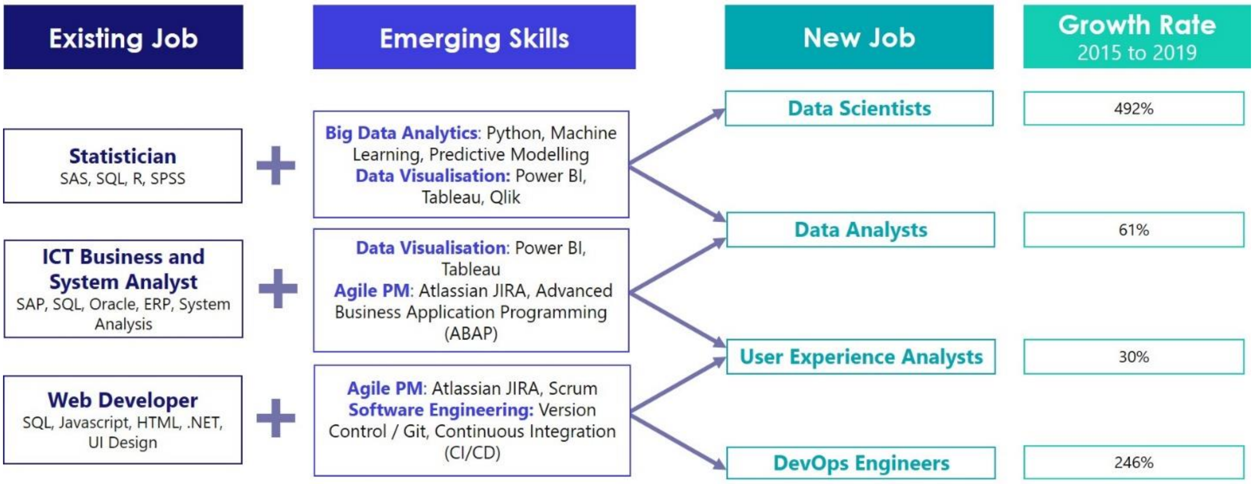
Additionally, the National Skills Commission used a mix of traditional sources of data and job posting data to identify 25 emerging occupations within the Australian labor market (National Skills Commission, 2020). By identifying emerging skills and looking at how these skills change existing jobs, they are able to recognize emerging or new jobs in the labor market. The emerging occupations are defined as new, frequently advertised jobs which are substantially different from occupations already defined in the Australian and New Zealand Standard Classification of Occupations (ANZSCO). The need to learn new skills accelerated during the COVID-19 pandemic, and using big data, the government could quickly analyze growing occupations into which they could funnel workers. The government used big data to understand quick changes in the labor market, complemented by traditional sources of data that helped them understand broader labor market trends. Figure 4.5 shows existing jobs paired with emerging skills, resulting in new and emerging occupations in the economy.

Figure 4.4 NERO data dashboard



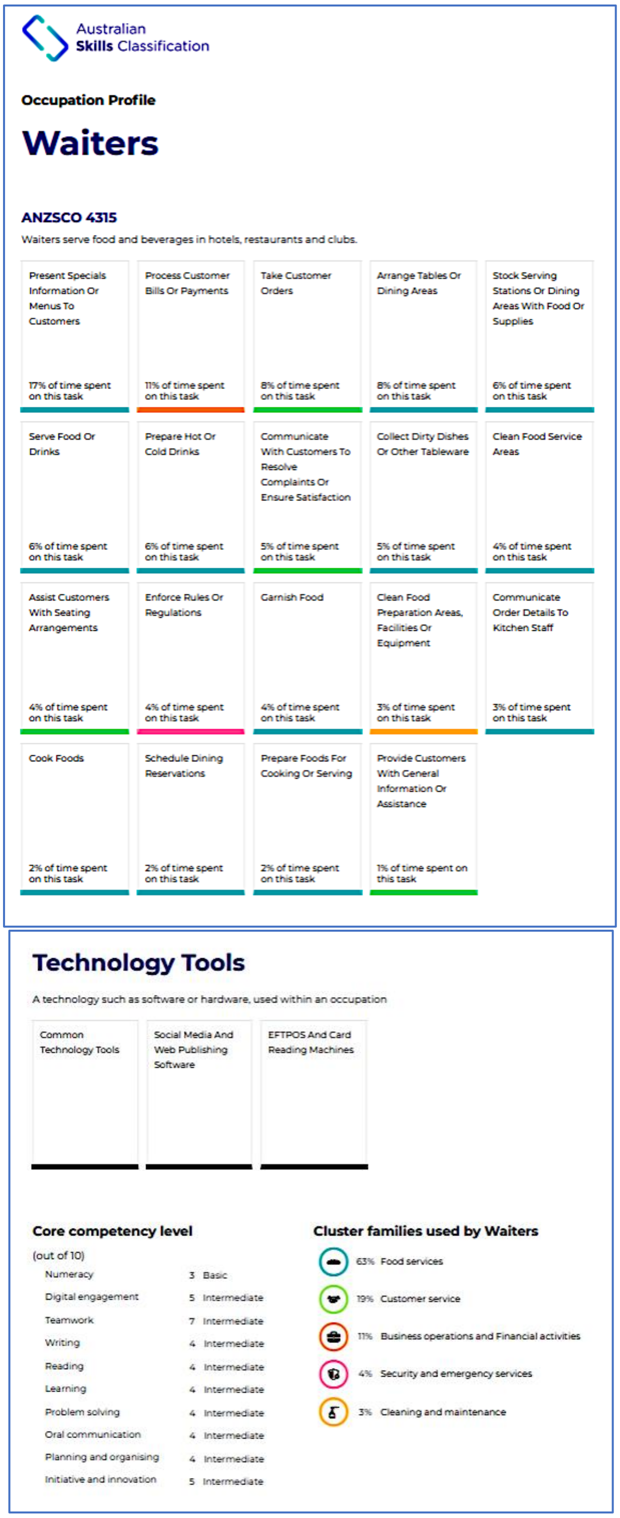
Source: National Skills Commission, Australia (n.d.-b).

Figure 4.5 Emerging occupations in Australia



Source: National Skills Commission, Australia (2020).

Figure 4.6 Australian Skills Classification (ASC) occupation profile for waiters



Source: National Skills Commission, Australia (n.d.-c).

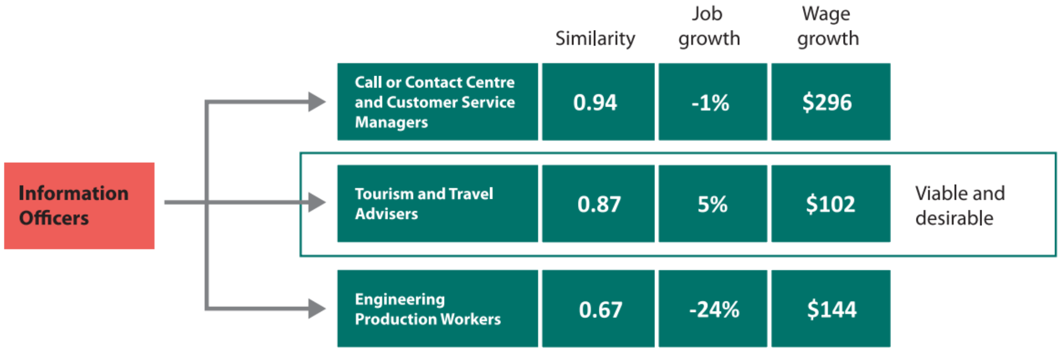
The National Skills Commission has also used job posting data to help validate the currency of skills in the Australian labor market for the March 2021 beta release of its Australian Skills Classification (ASC). The ASC complements the ANZSCO, providing a new level of detail about the skills that underpin Australian jobs. Intended to be a ‘common language’ for skills, the ASC provides ways to explore the connections and transferability of skills between occupations. The ASC will continue to be expanded and improved, and in conjunction with big data and other information sources, it will help identify changes within jobs, and new and emerging jobs, more readily. Figure 4.6 shows an example profile for waiters, and the interactive interface allows users to access additional information for each skill as well as navigate between connected skills and occupations.

By comparing the skills in one occupation to another, the ASC can be used to help create similarity measures between occupations. This can highlight potential job transition opportunities, and ASC data has been used to underpin Australian government reskilling tools such as Job Switch.

Reskilling workers

Another example of policymakers using big labor market data is in reskilling workers. The Department of Employment, Skills, Small and Family Business (now the Department of Education, Skills and Employment) partnered with the Boston Consulting Group and EBG to develop a report on reskilling Australia (Australian Government, 2019). The Australian government decided to use big labor market data to supplement the commonly used public skills data, O*NET, because the EBG data provides information that O*NET cannot, such as occupations coded to the local taxonomy (ANZSCO), education requirements coded to Australian degree systems (such as vocational education and training (VET), bachelor's degree or above), a dynamic and expansive skills taxonomy that includes emerging skills like data analytics and blockchain. An example of a career pathway for information officers that was developed using big labor market data is in Figure 4.7.

Figure 4.7 Job transitions for information officers in Australia



Source: Australian Government (2019, 24).

The Department of Education, Skills and Employment also used job posting data to publish weekly information on jobs in demand, by location, through the Jobs Hub app. This app was developed in response to the COVID-19 pandemic and the need to help people transition quickly to new roles during the labor market disruption. It allowed people to understand changes in their local labor market and how their existing skills could be applied to new roles.

Understanding emerging technologies and technological change

The Australian government worked to create two other tools using big labor market data. The Department of Industry, in conjunction with their agency AustCyber, uses big labor market data in a tool called CyberSeek (n.d.), which enables data exploration of the cybersecurity skills gap. The tool helps policymakers to answer questions about costs of hiring in their region, ease of hiring in their region, and whether they should source workers from other regions.

The government of Singapore also partners with big labor market data companies to inform their decisions. Singapore's Smart Nation and Digital Government Office (SNDGO) worked with LinkedIn Economic Graph as part of Singapore's AI Strategy. The goal of the collaboration was to explore where the necessary AI talent works and what skills they have across regions (Figure 4.8). Big labor market data enabled understanding of emerging skills and trends that is otherwise not possible with traditional data.

Figure 4.8 Top 5 originating economies for recently relocated AI-capable workers

Source: Tang et al. (2020).

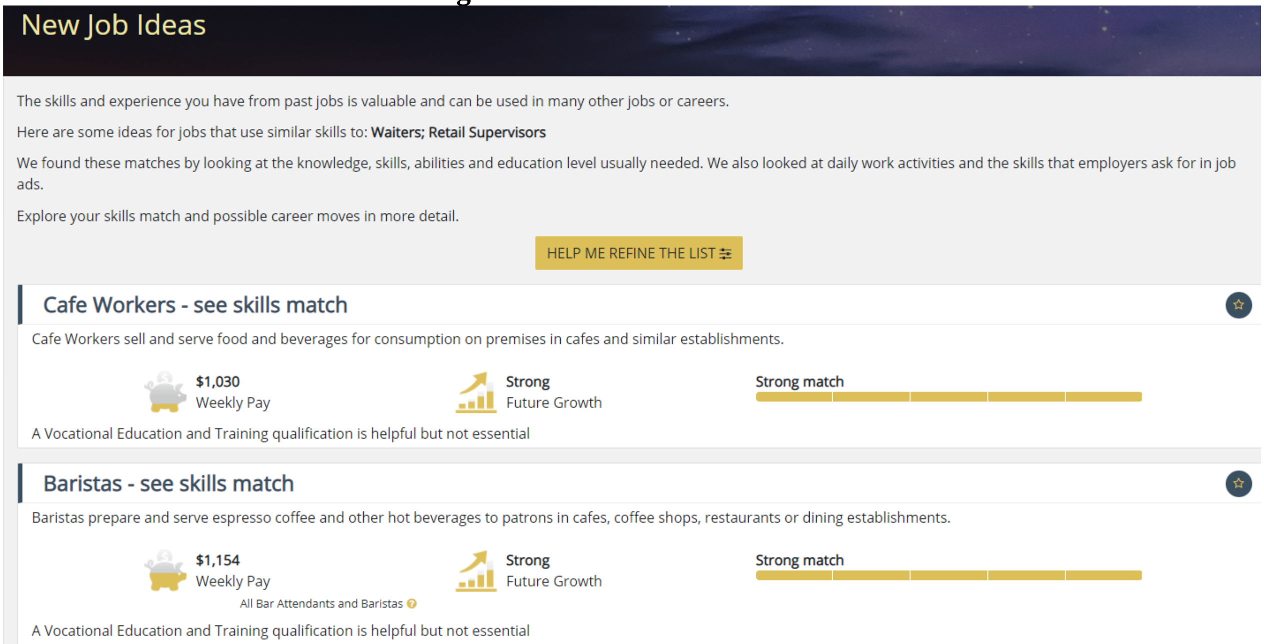
In addition, the Infocomm Media Development Authority of Singapore has been using EBG data since 2017 for early-stage analysis to help glean insights into local-talent demand to facilitate the making of policy decisions to help career counselors provide guidance around media and information and communications technology (ICT) roles.

Matching job seekers to employers and defining talent gaps and critical occupations

The European Union partners with labor analytics company EBG to develop Cedefop's online job vacancy analysis system. Cedefop, a European Union agency focused on the development of European VET policies, has been working with big labor market data since 2015 (European Commission, n.d.). The project is part of Eurostat's Big Data initiative, with this component focusing on web scraping job vacancies. Ultimately, the end goal is to create a pan-European online vacancy collection and analysis system that will provide data on the current and emerging skill needs of Europe and information relevant to policymaking. In partnership with EBG Europe (previously known as Tabulaex before acquisition), they conducted a feasibility study, identified the main job vacancy providers across European Union members, created a methodological approach to crawl, fetch and scrape, and collected job vacancy data.

Additionally, the Australian government developed a platform called Your Career that helps people find career information and publishes employment and education data (see Figure 4.9). The website allows users to look at occupations with current job vacancies based on a series of user input preferences.

Figure 4.9 Your Career tool



Source: Australian Government (n.d.).

The governments of Malaysia and Indonesia, in partnership with the World Bank, also worked with online job postings. The World Bank and the government of Malaysia released a report on Malaysia's skill shortages and critical occupations (World Bank, 2019). They use postings data to learn the skill and experience requirements of high-demand occupations and help create their list of critical occupations. This list is then used to align workforce development policies with employer demands. Similarly, Indonesia created a critical occupation list to highlight shortages and potentially strategic investment areas (World Bank, 2020). Additionally, they have an Online Skills and Vacancy Outlook initiative that collects online job postings by occupation (World Bank, 2021). These initiatives work to analyze skill imbalances and help policymakers to make investments in training programs and adjust incentives.

This type of skills gap analysis can also be used for equity purposes. In New Zealand, Tokona Te Raki is using big labor market data to drive longer-term systemic change to boost Māori success and tackle inequality. They aim to produce an understanding of current labor market data to support better employment outcomes for Māori and inform the business case for further investment in tools to enable future indigenous workforce development.

Digital and data policymakers

The Bureau of Economic Analysis in the US is developing an approach to using labor market data to measure the value of the data economy. Their approach uses job posting data from EBG to select data-related occupations based on skills and tasks, and then estimate labor costs of these occupations using Occupational and Employment Statistics (OES) data.

Similarly, the Measurement and Analysis of the Digital Economy (MADE) Group at the Organisation for Economic Co-operation and Development (OECD) has considered using big labor market data to understand technology adoption and measure the effect of ‘free’ digital products and services on welfare, which is largely not reflected in national accounts currently. There are limitations posed by survey data around technology adoption especially with open source software. Labor market data from job postings and social profiles can help to measure the adoption of open source technologies such as Python, TensorFlow and Perl, and the changes in demand over time as a share of all information technology (IT) jobs. By using a salary prediction model based on these technologies, value-added can be calculated.

Digital and data policymakers can also help encourage governments to standardize job posting and social profile data across economies, making it even easier for real-time labor market data to support policy analyses. In particular, helping to collect and standardize demand-side and supply-side labor market information would enable policymakers to capture with high degrees of accuracy economy-wide skill gaps and training needs. This could include:

- The standardization and collection of all job posting data in a database, including complete information such as salary ranges, core competencies, and degree and credential requirements.
- The support and production of a skills database of workers that includes profiles for all workers with credential attainments, certifications, skills, and employment histories.
- Cooperation across economies in taxonomy development and alignment, such as agreements to develop joint occupation and skills taxonomies or use international standard classification systems. This would allow for great cross-economy comparison and could help skill-migration policymakers to understand where to source talent from or where opportunities could be created for workers to be encouraged to support their local economies.

Macroeconomic policy analyses

There are a few ways in which macroeconomic policymakers can use big data to support their work. Economists at the Massachusetts Institute of Technology used job posting data to understand wage rigidity (Hazell et al., 2018). They found that posted wages change infrequently, wages for the typical job remained unchanged for 20 quarters, and that posted wages were especially unlikely to fall within a given job, implying downwards rigidity in the posted wage. They also found that posted wages were nearly acyclical for the typical job, suggesting substantial rigidity in the wage for new hires at the job level.

Big labor market data can be used to understand the effects of wage transparency, such as with Colorado’s new Pay Transparency Law. This law requires employers to (1) post compensation and benefits information for each job posting for Colorado jobs and (2) internally post promotional opportunities to current Colorado employees on the same day and sufficiently in advance of promotion decisions. Changes in wages based on these policy changes show up immediately in online job posting data, which enable the changes in wages or other job characteristics to be tracked.

GAINING INSIGHTS FROM BIG DATA

There are two main ways in which APEC economies can take the first step and gain insights into big data trends in their labor markets: (1) collaboration with academic researchers and (2) partnerships with non-governmental organizations (NGOs). While these modes do not allow for economies to tailor-fit their analyses and data collection processes, they can act as initial proof-of-concept opportunities and open doors for future collaboration or data integration.

Collaboration with researchers

Several examples of collaboration between academics and intermediaries of big labor market data have been discussed in Section 3 of this report. Academic institutions and research organizations can help provide necessary skills and analysis of available labor market information data at low or no cost to governments. Supporting researchers using new sources of data can accelerate the process by which economies gain a deeper understanding of new data and can assess the added value of new data relative to existing sources. Partnerships with research organizations and universities, including professors, PhD students and think tanks, can enable stakeholders to disseminate understanding of big data.

Researchers are also often heavy users of traditional labor market data, enabling synergies between current data sources and new, innovative sources. Academic research can assist in benchmarking datasets that currently do not have any benchmarking work done and can lend legitimacy to the use of big data analytics and new sources of data. Since the market for papers and articles in academic circles is rigorous, collaborating with researchers can accelerate the use of new data sources in wider circles.

Big data providers often allow use of their data for low or no cost for academic partnerships, making this option particularly cost-effective. Time investments are largely those of the research institution, and no additional training is required.

The key limitation of this approach is that researchers and academic circles have their own research interests, agendas and timelines. To the extent that these line up with stakeholder needs, this approach is sound, but it may not allow for quick insights or real-time updates in analysis.

An example of this type of partnership involves academics writing an OECD Working Paper to assess the AI-related jobs in Canada; Singapore; the UK; and the US (Squicciarini and Nachtigall, 2021). The researchers use EBG job posting data to identify AI jobs (defined as jobs that contain at least two AI skills).

Another example of this is a group of researchers using job posting data to compare Chinese and US demand for labor with respect to the impact of China’s five-year plans. The authors analyze the interdependence between the two economies and the impact of policy, using online postings data to ‘most directly measure firm’s desire to hire’ (Cen, 2021).

NGO partnerships

Partnerships between government stakeholders with multilateral organizations, such as APEC, and policy research-oriented NGOs like the World Economic Forum can help bring together efforts to incorporate new sources of labor market data into policy recommendations and streamline integration. Similar to academic researchers, these organizations often have in-house research and policy teams that are used to aggregating data from a wide variety of sources, including both traditional data and new big data. Additionally, these partnerships allow for a greater focus on cross-economy collaboration and policy, rather than purely academic research. The recommendations associated with work from NGO partnerships can help further APEC stakeholder needs and are more likely to be tailored to goals that APEC economies have.

As with academic partnerships, big data providers are more likely to provide data at low or no cost, especially if such partnerships are likely to increase their brand value and lend legitimacy to their data products. In terms of time and skill investment, since these multilateral organizations and NGOs take on the burden of analysis, this is minimal for government stakeholders. These partnerships are also likely to allow for increased collaboration and standardization of data use across economies, leading to economy-of-scale benefits that may not be available to individual entities.

An example of this type of arrangement was the partnership between APEC, LinkedIn and EBG (APEC 2020a). This partnership analyzed the digital skills gap to provide guidance to APEC economies in navigating the growing digitalization amid the COVID-19 pandemic. It allowed the team to develop APEC-wide actions to close the digital skills gap, creating cross-economy collaboration and next steps.

Another example of this type of partnership is the World Bank Group–LinkedIn partnership (Zhu et al., 2018). This partnership was designed to use LinkedIn data to inform policy across 100+ economies. The partnership worked together to define LinkedIn data characteristics, then LinkedIn extracted and validated the data, and from there the data was used to provide results by economy within the World Bank Group.

COLLECTING BIG DATA AND COLLABORATING WITH BIG DATA PROVIDERS

This section outlines the step-by-step process to start integrating big data into policymaking by collaborating with big data providers.

Partner with intermediaries to minimize limitations

Governments can directly collect their own real-time labor market data and incorporate it into labor market information systems. This can include matching of job postings, social profiles or online skill databases with educational and qualification data. Merging supply-side and demand-side labor market data could help connect workers, students and job seekers with employers, and those seeking training with appropriate education opportunities to advance their career pathways. This would provide the most opportunity to align supply and demand within an economy and across the APEC region.

The primary limiting factors to this process are costs in terms of financial investment, time, and required skills embedded in agencies. Recreating processes to analyze and systematize millions of rows of data, including data engineering and parsing requirements, would entail substantial investment. Risks to this method also include the requirement for long-term maintenance. While up-front costs would likely be highest (in terms of beginning to process the data), maintenance and improvements would be required. This approach also may create silos across APEC economies, especially if each economy works to reinvent the wheel and approaches this issue from different angles and with different technologies.

In order to minimize risks from these potential limitations, we recommend that APEC economies work with a private intermediary who has already developed the capacity to collect, organize and analyze the data. Since many new labor market data sources are part of private companies' data collection efforts, public–private partnerships can help strengthen data pipelines and provide member economies with direct access to real-time data. APEC economies can purchase access directly to data feeds or tools, including through product services, APIs, flat files or other methods. This partnership can allow for extremely tailored analyses for each APEC economy, depending on desire and capacity.

Data formats and access

Partnering with entities that offer their data in product or API forms is likely the most straightforward way to access and display the data. This enables economies to have a more plug-and-play option without requiring extensive data infrastructure.

Alternatively, these partnerships can also include direct feeds of data that stakeholders can opt to analyze and process themselves. This may be useful if product options do not allow for the level of granularity or customization desired. For this, APEC economies could consider additionally partnering with researchers at intermediary institutions or with academics or others who have familiarity with the data. This can help smooth onboarding processes and provide best practices for issues such as loading and processing the data, storing the data and analyzing the data.

Researchers at partner international organizations may be willing to lend technical expertise to these partnerships, allowing for new data collection to begin and for benefits to be realized to both the economy and intermediaries looking to advance their data collection. An example of this partnership is collaboration between The World Bank, the government of Malaysia, and EBG in their report on Malaysia's skill shortages and critical occupations (World Bank, 2019). The World Bank uses postings data to help create their list of critical occupations by using online job posting data as a proxy for job openings and to learn the skill and experience requirements of high-demand occupations.

Table 4.1 outlines some of the main challenges APEC governments may face and the recommended approaches and solutions to overcome these challenges.

NEXT STEPS FOR POLICYMAKERS

The next steps can be to work with a third-party intermediary to help begin collecting the data:

1. **Understand and assess the full scale of costs.** In any of the partnership modes described above, there are up-front costs as well as dynamic costs. It is important to consider the costs of initial creation and set-up, people and analytical resources, and maintenance of taxonomical and data updates.

2. **Begin with a small-scale pilot project**. This pilot project should aim to solve a very specific problem, like determining top skills required for each occupation in the economy. Initially, the government may face some hesitation from people exclusively used to traditional data. It is important that the first project starts small to gain trust in the data.

3. **Once trust is gained, start a larger project.** After the initial data scoping to ensure big labor market data can be helpful and used to solve a problem, the government can consider larger scale projects. Further projects can include work like the examples described earlier, including identifying labor shortages, emerging skills or growing occupations.

Table 4.1 Challenges and Recommendations for APEC governments

Challenges	Recommendations
High up-front costs to collect online labor market data, including: • Development of technical capabilities to automatically scrape data and identify online text as job postings, résumés, etc. • Storage space for storing large raw text data • Finding the best sources of job postings or other online labor market data, including job boards and individual employer website	• Work with a third-party intermediary familiar with data collection in other geographies to build aggregation system for job postings • Use local knowledge of labor market to identify high-density sources of labor market data, such as the economy's primary job board
Cleaning, deduplicating, and preparing raw text data for analysis requires advanced modeling expertise including: • Technical capability to see a job posting or social profile/résumé on multiple websites and deduplicate across sources • Ability to read raw text in the source language and categorize the data into labeled fields, such as employer, occupation, location, skills, etc.	• Apply a parser from a third-party intermediary built on at least one local language as a pilot case • Limit sources to one large job board or one set of employers to ensure the parsing of postings can be more readily standardized • Work with local language experts or translators to translate some components of a parser into a local language
Merging data with current government sources and taxonomies and analyzing big labor market data • Big labor market data may not be representative of the population and may be limited to more urban, digital or high-skill jobs. This may be a particular challenge for developing economies. • Maintaining a database of big labor market data requires near-constant updating, and ever-growing storage capacity	• Utilize crosswalks and parsers that parse to both local taxonomies and international standard taxonomies, which third-party intermediaries often have • Match the distribution of big labor market data to public datasets to understand gaps and differences • Plan for expanding storage costs as improvements in technology mean that an additional number of job postings/résumés are collected monthly or annually
Visualizing data and creating models and tools on top of big labor market data to answer research and policy questions • Dashboards and interfaces make real-time data accessible to a wider range of agencies rather than just technical experts, but need to be built and maintained • Big data analysis requires a different understanding of data limitations, such as changes in data collection over time affecting the time series analysis	• Build a tool in partnership with a product organization to easily view different cuts of data, such as top-growing skills or top industries

5. CONCLUSION

As digitalization occurs across all corners of the economy, big data will continue to expand and increase in value. While traditional labor market data is collected by most APEC economies and continues to be an important source of information, these data leave gaps that are important to consider. Most traditional sources of data are collected through long, costly and time-intensive surveys that paint a broad picture of the labor market but allow limited strategic policymaking. Aggregate statistics mask nuances within certain sectors, occupations or regions that could make a difference in approaches to reskilling or upskilling, social protection policies, and education and training provision.

The timeliness and breadth of big data offer much-needed detail and an opportunity for policymakers to react in real time. Big data also provides an additional layer of granularity, helping stakeholders to pinpoint areas of growth and decline, and craft appropriate policy responses. The ways in which people interact in labor markets are also changing: increasing trends in remote work, online communication technologies and general switching of tasks from in-person to online means that it is harder to track the nature of work through traditional survey methods. Data from online collaboration tools and other tracking systems can help provide more context for how workers operate and the skills required to do so successfully.

The shift to an increasingly digital world means that more and more data are being collected and utilized by private corporations in a way that was not possible in previous eras. Data are used to target advertising and marketing campaigns, hire and recruit talent, and make strategic decisions. Without incorporating elements of big data analysis and collection into public sector processes, APEC governments run the risk of falling behind. The efficacy of wide ranges of government programs, including training, social service provision, and other public goods could be substantially increased with the use of new sources of big data.

While it is true that there is significant cost associated with adopting big data techniques, including in terms of collection and data engineering as well as hiring skilled talent, there are several ways to mitigate these challenges and benefit from the advantages offered by big data. Partnering with academics and universities can decrease the start-up costs to using this data and provide useful benchmarking and comparisons to already-known public datasets. Collaboration between the public sector and non-governmental organizations (NGOs) or private entities directly can also speed up the time it takes to ingest and use big data. Certain governments have already taken advantage of these partnership opportunities and are starting to incorporate some elements of big data alongside their traditional labor market information systems. This has allowed them to understand critical occupations, skills and sectors and create targeted policies to support a productive economy.

The challenges brought by the COVID-19 pandemic were immediate and immense. The lack of real-time, granular data across labor markets, health outcomes and other important economic indicators constrained economies' ability to develop and adapt policy in a timely and targeted manner. APEC economies can start to use big data now to prepare for the next economic shock and more readily respond to any future volatility. While not a silver bullet, big labor market data provides valuable complementarities to traditional labor market data and can be a cost-effective way of supporting economic growth, workers, students and employers.

APPENDIX

Sources used in Table 2.1 and Table 2.2:

APEC member economy	Sources
Australia	- Australian Bureau of Statistics. "Labour Force, Australia, Detailed." 27 May 2021, https://www.abs.gov.au/statistics/labour/employment-and-unemployment/labour-force-australia-detailed/apr-2021. Accessed 30 June 2021. - Australian Bureau of Statistics. "Labour Force, Australia Methodology.", https://www.abs.gov.au/methodologies/labour-force-australia-methodology/may-2021. Accessed 30 June 2021.
Brunei Darussalam	- Department of Economic Planning and Statistics - Brunei Darussalam. http://www.deps.gov.bn/SitePages/Labour%20Force.aspx. Accessed 30 June 2021. - "2020 APEC Economic Policy Report." Asia-Pacific Economic Cooperation, Nov. 2020, https://www.apec.org/Publications/2020/11/2020-APEC-Economic-Policy-Report. Accessed 30 June 2021.
Canada	- "Labour Force Survey (LFS)." Statistics Canada, https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3701. Accessed 8 October 2021. - "National Occupational Classification." Government of Canada, https://noc.esdc.gc.ca/. Accessed 29 July 2021. - "North American Industry Classification System (NAICS) Canada 2012." Statistics Canada, 11 Jan. 2019, https://www.statcan.gc.ca/en/subjects/standard/naics/2012/index. - "Statistics Canada." Labor, 30 June 2021, https://www150.statcan.gc.ca/n1/en/subjects/Labour. Accessed 30 June 2021.
Chile	- Employment Survey. Statistical Institute (Chile), Apr. 2010, https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/english/methodology.pdf?sfvrsn=70e6e74d_5. Accessed 30 June 2021. - SEPARATA TÉCNICA ANUAL. Statistical Institute, Jan. 2020, https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/anuales/separata-2019.pdf?sfvrsn=4a7591c2_4. Accessed 30 June 2021. - Methodological Document Employment Survey. Statistical Institute, Feb. 2020, https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/english/methodology-2020.pdf?sfvrsn=68b2ca2b_4. Accessed 30 June 2021.
People's Republic of China	- "China Statistical Yearbook 2019." Bureau of Statistics of China, http://www.stats.gov.cn/tjsj/ndsj/2019/indexeh.htm. Accessed 30 June 2021. - Cook, Sarah, and James Keeley. Micro-Data Scoping Study - China. Jan. 2007. Zotero, https://esrc.ukri.org/files/about-us/policies-and-standards/national-data-strategy/international-microdata-scoping-studies-project-china/. Accessed 30 June 2021.
Hong Kong, China	- "Labor Force, Employment and Unemployment." Census and Statistics Department, https://www.censtatd.gov.hk/en/scode200.html#section3. Accessed 30 June 2021.
Indonesia	- Labor Force Situation in Indonesia. Statistics Indonesia, Feb. 2015, https://www.ilo.org/public/libdoc//igo/P/76333/76333(feb2015)267.pdf. Accessed 30 June 2021.
Japan	- "Historical Data." Statistics Bureau of Japan, https://www.stat.go.jp/english/data/roudou/lngindex.html#note. Accessed 30 June 2021. - "List of Datasets." E-Stat: Portal Site of Official Statistics of Japan, https://www.e-stat.go.jp/en/stat-search/database?page=1&layout=datalist&toukei=00200531&kikan=00200&tstat=000000110001&cycle=2&tclass1=000001040286&tclass2=000001040298&tclass3=000001038122&result_page=1&tclass4val=0. Accessed 30 June 2021.
Republic of Korea	- "Statistical Database." KOSIS Korean Statistical Information Service, https://kosis.kr/eng/statisticsList/statisticsListIndex.do?menuId=M_01_01&vwcd=MT_ETITLE&parmTabId=M_01_01. Accessed 30 June 2021.
Malaysia	- "Key Statistics of Labor Force in Malaysia, April 2021." Department of Statistics Malaysia Official Portal, June 2021, https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=124&bul_id=UXQ2VUpIOVRvT3p6T0NHMHk1TGkxUT09&menu_id=Tm8zcnRjdVRNWWlpWjRlbmtlaDk1UT09. Accessed 30 June 2021.

Mexico	- Cómo se hace la ENOE. Métodos y procedimientos. Instituto de Estadística y Geografía, 2007, https://en.www.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825190613.pdf. Accessed 1 July 2021.
New Zealand	- “Labor Market Statistics: March 2021 Quarter.” Stats NZ, May 2021, https://www.stats.govt.nz/information-releases/labour-market-statistics-march-2021-quarter. Accessed 1 July 2021.
Papua New Guinea	- “2020 APEC Economic Policy Report.” Asia-Pacific Economic Cooperation, Nov. 2020, https://www.apec.org/Publications/2020/11/2020-APEC-Economic-Policy-Report. Accessed 30 June 2021. - Labor Mobility and Labor Market Data: A Baseline Study of APEC Economies. International Labor Organization (ILO), 2019, https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-bangkok/documents/publication/wcms_737366.pdf. Accessed 1 July 2021.
Peru	- Labor Mobility and Labor Market Data: A Baseline Study of APEC Economies. International Labor Organization (ILO), 2019, https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-bangkok/documents/publication/wcms_737366.pdf. Accessed 1 July 2021.
The Philippines	- “Classification Systems.” Philippine Statistics Authority OpenSTAT, https://openstat.psa.gov.ph/Metadata/Classification-Systems. Accessed 1 July 2021.
Russia	- “Labor Market: Summary Methodology.” Department of Labor Statistics, https://www.gks.ru/bgd/free/b00_25/IssWWW.exe/Stg/d000/I000790R.HTM. Accessed 1 July 2021.
Singapore	- “Special Data Dissemination Standard.” International Monetary Fund, 2021, https://dsbb.imf.org/sdds/dqaf-base/country/SGP/category/EMP00. Accessed 1 July 2021. - “Singapore Standard Occupational Classification SSOC 2020.” Department of Statistics Singapore, https://www.singstat.gov.sg/standards/standards-and-classifications/ssoc. Accessed 30 March 2021. - “Singapore Standard Industrial Classification SSIC 2020.” Department of Statistics Singapore, https://singstat.gov.sg/standards/standards-and-classifications/ssic. Accessed 30 March 2021.
Chinese Taipei	- General Description of Statistical Methods. Zotero, https://eng.stat.gov.tw/public/Data/912301123140ZT9Z70G.pdf. Accessed 1 July 2021.
Thailand	- “Labor Branch.” Statistical Office (Thailand), http://statbbi.nso.go.th/staticreport/page/sector/en/02.aspx. Accessed 1 July 2021. - “Labor Force Survey.” Statistical Office, http://web.nso.go.th/eng/stat/lfs/lfse.htm. Accessed 1 July 2021.
United States	- “Current Employment Statistics - CES (Federal).” US Bureau of Labor Statistics, https://www.bls.gov/ces/data/employment-and-earnings/. Accessed 1 July 2021. - “Employment Projections.” US Bureau of Labor Statistics, https://www.bls.gov/emp/. Accessed 1 July 2021. - “BLS Restricted Data Access.” US Bureau of Labor Statistics, https://www.bls.gov/rda/restricted-data.htm#employment. Accessed 1 July 2021.
Viet Nam	- Report on Labor Force Survey 2019. General Statistics Office, 2019, https://www.gso.gov.vn/wp-content/uploads/2021/05/labor-force-report-2019.pdf. Accessed 1 July 2021. - Demombynes, Gabriel, and Mauro Testaverde. Employment Structure and Returns to Skill in Vietnam: Estimates Using the Labor Force Survey. World Bank, Washington, DC, Mar. 2018. DOI.org (Crossref), doi:10.1596/1813-9450-8364. Accessed 1 July 2021.
Overall	- “ILOSTAT.” ILOSTAT, https://ilostat.ilo.org/. Accessed 30 June 2021. - “Country Profiles.” United Nations Statistics Division, https://unstats.un.org/unsd/dnss/cp/searchcp.aspx. Accessed 1 July 2021.

Additional methodological details for Table 2.3:

Economy	**Methodological details**
Australia	Households within selected homes interviewed every month for 8 months - 1/8 sample replaced each month - Sample includes 8 sub-samples (rotation groups) where each sample is in the survey for 8 months - A new rotation group introduced each month, replacing another rotation group (usually from same geographic area) - Interviews conducted by trained interviewers or through online self-survey - First interview is usually face-to-face and next interviews are over the phone
Chile	- Primarily sampling units constructed from 2017 Census and 2016 Precensus o Primarily sampling units correspond to geographical areas that are homogenous in number of frame dwellings (private dwellings recorded in 2017 Census) o Average size of urban primarily sampling unit is 200 dwellings and average size of rural primarily sampling unit is 90 dwellings - Secondary sampling unit is a private occupied dwelling within the selected primary sampling unit - Selection of primary sampling units is done with systematic random selection with probability proportional to size - Homes within each selected primary sampling unit are selected independently and systematically. Each home has an equal probability of being selected - Panel survey - Individuals interviewed more than once and at different times to measure change in indicators over time
Republic of Korea	- Sample selected based on results of every year's Population and Household Census - Sampling unit replaced by 1/36 of rotation sampling every month - Survey conducted by computer-assisted personal interviewing (CAPI), computer assisted self-interviewing (CASI),
The Philippines	- Randomly assigned and selected set of geographic areas with non-overlapping boundaries known as primary sampling units o Primary sampling unit can be the whole *barangay* (smallest political subdivision in the economy), a portion of a large barangay, or combinations of small barangays o 3,416 primary sampling units - 117 major domains: 81 provinces, 33 urbanized cities and 3 other areas o Each domain is divided into primary sampling units with 100–400 households - Average of 12 sample households for highly urbanized cities and average of 16 sample households for primary sampling units in provinces - Computer-aided personal interviewing using tablets
Thailand	- Stratified two-stage sampling - Provinces constituting a strata – 77 altogether - Primary sampling units were enumeration areas for municipal areas and non-municipal areas and private households o Selection was separate and independent - Secondary sampling units were private households/persons in collective households - Information obtained through face-to-face interviews with heads of members of households

REFERENCES

Altamirano, A. and N. Armaral. 2020. “A Skills Taxonomy for LAC: Lessons Learned and a Roadmap for Future Users” Inter-American Development Bank. https://publications.iadb.org/publications/english/document/A-Skills-Taxonomy-for-LAC-Lessons-Learned-and-a-Roadmap-for-Future-Users.pdf.

Ammanath, B., S. Hupfer, and D. Jarvis. 2020. *Thriving in the Era of Pervasive AI: Deloitte’s State of AI in the Enterprise, 3rd edn*. Deloitte. https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/state-of-ai-and-intelligent-automation-in-business-survey.html

APEC. 2017. “2017 APEC Economic Policy Report.” Singapore: APEC. https://www.apec.org/publications/2017/11/2017-apec-economic-policy-report

———. 2020a. “APEC Closing the Digital Skills Gap Report: Trends and Insights.” Singapore: APEC. https://www.apec.org/Publications/2020/12/APEC-Closing-the-Digital-Skills-Gap-Report.

———. 2020b. “2020 APEC Economic Policy Report.” Singapore: APEC. https://www.apec.org/Publications/2020/11/2020-APEC-Economic-Policy-Report

———. 2021c. “2021 APEC Economic Policy Report.” Singapore: APEC. https://www.apec.org/publications/2021/11/2021-apec-economic-policy-report

Australian Government. 2019. “Reskilling Australia: A Data-Driven Approach.” Department of Employment, Skills, Small and Family Business. https://www.dese.gov.au/employment-research-and-statistics/resources/reskilling-australia-data-driven-approach.

———. n.d. “Your Career.” Accessed 3 August 2021 https://yourcareer.gov.au/about-us/

Automatic Data Processing (ADP). 2021. “2021 Corporate Overview.” May 2021, https://www.adp.com/~/media/Corporate%20Overview/ADP-Corporate-Overview.ashx

Balakrishnan, T., M. Chui, B. Hall et al. 2020. “The State of AI in 2020.” McKinsey, 17 November 2020. “https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020#

Bartik, A.W., M. Bertrand, F. Lin et al. 2020. “Measuring the Labor Market at the Onset of the COVID-19 Crisis.” Working Paper, National Bureau of Economic Research.

Belot, Michèle, Philipp Kircher, and Paul Muller. 2019. “Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice.” *The Review of Economic Studies* 86 (4): 1411–47. https://doi.org/10.1093/restud/rdy059.

Booth, K., J. Carrese, L. Coleman et al. 2015. “Understanding Labor Market Information Resources: Descriptions, Benefits, and Limitations.” Centers of Excellence for Labor Market Research. http://coeccc.net/COE/media/SupplyandDemandPageDocuments/Understanding-Labor-Market-Information-Resources.pdf

Cajner, T., L.D. Crane, R.A. Decker et al. 2019. “Improving the Accuracy of Economic Measurement with Multiple Data Sources: The Case of Payroll Employment Data.” Working Paper, National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w26033/w26033.pdf.

Cajner, T., L.D. Crane, R.A. Decker et al. 2020. “Tracking Labor Market Developments during the COVID-19 Pandemic: A Preliminary Assessment.” Finance and Economics Discussion Series 2020, no. 030. https://doi.org/10.17016/FEDS.2020.030.

Cammeraat, E., and M. Squicciarini. 2021. *Burning Glass Technologies' Data Use in Policy-Relevant Analysis: An Occupation-Level Assessment*. OECD. https://doi.org/10.1787/cd75c3e7-en.

Carnevale, A.P., T. Jayasundera, and D. Repnikov. 2014. "Understanding Online Job Ads Data: A Technical Report." Washington, DC: Center on Education and the Workforce, McCourt School of Public Policy, Georgetown University. https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM.Tech_.Web_.pdf

Cen, X. 2021. "A Race to Lead: How Chinese Government Interventions Shape the US-China Production Competition." SSRN Scholarly Paper 3564494, 25 May 2021. https://doi.org/10.2139/ssrn.3564494.

Chetty, R., J. Friedman, N. Hendren et al. 2020. *The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data*. Working Paper 27431, National Bureau of Economic Research. https://doi.org/10.3386/w27431

Cho, D. 2018. *The Labor Market Effects of Demand Shocks: Firm-Level Evidence from the Recovery Act*. Princeton University and ADP.

City & Guilds Group and Burning Glass Technologies. 2021. "Building Bridges towards Future Jobs." https://www.burning-glass.com/wp-content/uploads/2021/02/CG_Skills-gap-report_A4_34pp_Final_Single-pages.pdf

City of New York. n.d. "NYC Analytics: Recovery Data Partnership." Accessed 16 July 2021. https://www1.nyc.gov/site/analytics/initiatives/recovery-data-partnership.page

Cognilytica. 2020. "Global AI Adoption Trends & Forecast 2020." 22 January. https://www.cognilytica.com/2020/01/22/global-ai-adoption-trends-forecast-2020/

CompTIA. 2017. "Assessing the IT Skills Gap." Research report, May. https://comptiacdn.azureedge.net/webcontent/docs/default-source/research-reports/research-report---comptia-it-skills-gap.pdf?sfvrsn=99cb420e_0.

Cook, S. and J. Keeley. 2007. "Micro-data Scoping Study – China." https://esrc.ukri.org/files/about-us/policies-and-standards/national-data-strategy/international-microdata-scoping-studies-project-china/

CyberSeek. n.d. Website. Accessed 2 August 2021. https://www.cyberseek.com.au/

Davenport, T.H. and D. J. Patil. 2012. "Data Scientist: The Sexiest Job of the 21st Century." *Harvard Business Review*, October. https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century

Demombynes, G. and M. Testaverde. 2018. *Employment Structure and Returns to Skill in Vietnam: Estimates Using the Labor Force Survey*. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-8364

Dickerson, A., R. Wilson, G. Kik et al. 2012. "*Developing Occupational Skills Profiles for the UK: A Feasibility Study*." Evidence Report 44, UK Commission for Employment and Skills.

European Commission. 2017. *ESCO Handbook: European Skills, Competences, Qualifications and Occupations, 1st edn.* European Commission, Directorate General for Employment, Social Affairs and Inclusion. https://data.europa.eu/doi/10.2767/934956

———. n.d. "WPB Online Job Vacancies." European Commission: CROS. Accessed 4 September 2021. https://ec.europa.eu/eurostat/cros/content/WPB_Online_job_vacancies_en

Faberman, R.J. and M. Kudlyak. 2016. "What Does Online Job Search Tell Us about the Labor Market?" *Economic Perspectives* (2016): 15.

Forsythe, E., L.B. Kahn, F. Lange et al. 2020. "Labor Demand in the Time of COVID-19: Evidence from Vacancy Postings and UI Claims." *Journal of Public Economics* 189 (September): 104238. https://doi.org/10.1016/j.jpubeco.2020.104238

Hazell, J., M. Kazemi, and B. Taska. 2018. "Posted Wage Rigidity."

Horton, J.J. and P. Tambe. 2015. "Labor Economists Get Their Microscope: Big Data and Labor Market Analysis." *Big Data* 3, no. 3 (September): 130–37. https://doi.org/10.1089/big.2015.0017

ICIMS. n.d. "Recruiting & Talent Acquisition Solutions | ICIMS." Accessed 19 July 2021. https://www.icims.com/solutions/

ILOSTAT, n.d. Website. International Labour Organization. Accessed 30 June 2021. https://ilostat.ilo.org/.

International Labour Organization (ILO). 2019. "Labour Mobility and Labour Market Data: A Baseline Study of APEC Economies." Report prepared for the *APEC Labour Mobility Statistics Forum*, Viña del Mar, Chile, 2–3 May. https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-bangkok/documents/publication/wcms_737366.pdf

Jyoti, R., and M. Shirer. 2020. "IDC Survey Finds Artificial Intelligence Adoption Being Driven by Improved Customer Experience, Greater Employee Efficiency, and Accelerated Innovation." IDC, 10 June 2020. https://www.idc.com/getdoc.jsp?containerId=prUS46534820

KPMG. *Thriving in an AI World*. KPMG, April 2021. https://info.kpmg.us/content/dam/info/en/news-perspectives/pdf/2021/Updated%204.15.21%20-%20Thriving%20in%20an%20AI%20world.pdf.

Kronos. "Kronos Has Record 2019: Annual Recurring Revenue Crosses $1 Billion." 30 October 2019. https://www.kronos.com/about-us/newsroom/kronos-announces-record-2019-annual-recurring-revenue-crosses-1-billion.

LinkedIn. 2019. "AI Talent in the European Labour Market." LinkedIn Economic Graph, November. https://economicgraph.linkedin.com/content/dam/me/economicgraph/en-us/reference-cards/research/2019/AI_Talent_in_the_European_Labour_Market_executive_summary.pdf.

———. n.d. "LinkedIn Career Explorer." Accessed 20 July 2021. https://linkedin.github.io/career-explorer/

Loukides, M. 2021. "O'Reilly – AI Adoption in the Enterprise 2021." O'Reilly. https://get.oreilly.com/ind_ai-adoption-in-the-enterprise-2021.html

Mariani, M. 1999. "*Replace with a Database: O*NET Replaces the Dictionary of Occupational Titles*." Occupational Outlook Quarterly (Spring): 3–9.

Markow, W. and R. Sederberg. 2020. "*Skills of Mass Disruption: Pinpointing the 10 Most Disruptive Skills in Tech*." December. https://www.burning-glass.com/wp-content/uploads/2020/12/Skills-of-Mass-Disruption-Report.pdf.

National Skills Commission, Australia. 2020. "Emerging Occupations." 24 August. https://www.nationalskillscommission.gov.au/sites/default/files/2020-08/Emerging%20occuptations_How%20new%20skills%20are%20changing%20Australian%20jobs_0.pdf

———. n.d.-a. "Jobs and Education Data Infrastructure (JEDI)." Accessed 2 August 2021. https://www.nationalskillscommission.gov.au/our-work/jobs-and-education-data-infrastructure-jedi

———. n.d.-b. "Nowcast of Employment by Region and Occupation." Accessed 2 August 2021. https://www.nationalskillscommission.gov.au/our-work/nero

———. n.d.-c. "Australian Skills Classification." Accessed 2 August 2021. https://www.nationalskillscommission.gov.au/our-work/australian-skills-classification

O'Kane, L., R. Narasimhan, J. Nania et al. 2020. "Digitalization in the German Labor Market: Analyzing Demand for Digital Skills in Job Vacancies." Gütersloh: Bertelsmann Stiftung. https://www.bertelsmann-stiftung.de/fileadmin/files/user_upload/STUDIE_Burning_Glass_EN_FINAL.pdf

O*NET. 2017. "PDF Questionnaires." O*NET Data Collection Program. Last updated 26 April 2017. https://onet.rti.org/pdf/

———. n.d.-a. "Summary Report for 41-4011.00 - Sales Representatives, Wholesale and Manufacturing, Technical and Scientific Products." O*NET OnLine. Accessed 20 July 2021. https://www.onetonline.org/link/summary/41-4011.00#Skills

———. n.d.-b. "The O*NET® Content Model." O*NET Resource Center. Accessed 20 July 2021. https://www.onetcenter.org/content.html#cm1

Opportunity Insights. n.d. "Economic Tracker." accessed 10 September 2021. www.tracktherecovery.org.

PwC. 2019. "CEOs Focus on Upskilling Their Workforce to Close the Skills Gap and Boost Innovation and Trust, Says PwC." Media release, PwC, 15 July. https://www.pwc.com/gx/en/news-room/press-releases/2019/ceos-upskilling-workforce.html

Reeder, M.C., T.C. Burgoyne, and M.T. Allen. 2020. "*O*NET® Analyst Ratings of Occupational Skills: Analysis Cycle 21 Results*." October. https://www.onetcenter.org/dl_files/AOSkills_21.pdf

Society for Human Resource Management (SHRM). 2019. "Skills Shortage Tightens Job Market; 83% of HR Professionals Report Difficulty Recruiting." Media release, 5 February. https://www.shrm.org/about-shrm/press-room/press-releases/pages/skills-gap-research-workplace-immigration-report.aspx

Squicciarini, M. and H. Nachtigall. 2021. "Demand for AI Skills in Jobs: Evidence from Online Job Postings." Working paper, OECD. https://doi.org/10.1787/3ed32d94-en

Statista. "Public Cloud Software Services (SaaS) Penetration Rate Worldwide by Application 2015-2020." n.d. Accessed November 29, 2021. https://www.statista.com/statistics/782240/worldwide-software-as-a-service-applications-penetration-rate/.

Tang, W., P.Y. Chua, and A. Tan. 2020. "Assessing Singapore's AI-Capable Workers." LinkedIn Economic Graph, 20 March 2020. https://economicgraph.linkedin.com/blog/assessing-singapore-s-ai-capable-workers

Turrell, A., B.J. Speigner, J. Djumalieva et al. 2019. "*Transforming Naturally Occurring Text Data into Economic Statistics: The Case of Online Job Vacancy Postings*." Working Paper 25837, National Bureau of Economic Research. https://doi.org/10.3386/w25837.

United Nations Statistics Division. n.d.-a. "Country Profiles," accessed 1 July 2021, https://unstats.un.org/unsd/dnss/cp/searchcp.aspx

———. n.d.-b. "Country Profile of Indonesia." Accessed 1 July 2021. https://unstats.un.org/unsd/dnss/cp/searchcp.aspx

Vilvovsky, N., M. Walsh, and W. Markow. 2020. "Skilling Up for the Future: The Growing Demand for Salesforce Talent." Burning Glass Technologies. https://developer.salesforce.com/resources2/career2017/docs/Skilling+Up+for+the+Future.pdf.

Vilvovsky, N., M. Walsh, W. Markow et al. 2021. "Visualizing the Future Globally: Tracking Worldwide Demand for 3D Graphics Skills." Burning Glass Technologies. https://www.burning-glass.com/wp-content/uploads/2021/01/Visualizing-the-Future-Globally-1.pdf

World Economic Forum (WEF). 2018a. *The Future of Jobs Report 2018*. Geneva: WEF Centre for the New Economy and Society. http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf

———. 2018b. "Towards a Reskilling Revolution: A Future of Jobs for All." http://www3.weforum.org/docs/WEF_FOW_Reskilling_Revolution.pdf.

———. 2019. "Towards a Reskilling Revolution: Industry-led Action for the Future of Work." WEF, Centre for New Economy and Society. http://www3.weforum.org/docs/WEF_Towards_a_Reskilling_Revolution.pdf.

World Bank. 2019. *Monitoring Occupational Strategies*. World Bank. https://documents1.worldbank.org/curated/en/877271568138256963/pdf/Monitoring-Occupational-Shortages-Lessons-from-Malaysias-Critical-Occupations-List.pdf

———. 2020. "Indonesia's Critical Occupations List 2018." https://documents1.worldbank.org/curated/en/763611585857010121/pdf/Indonesias-Critical-Occupations-List-2018-Technical-Report.pdf

———. 2021. "Toward a World-Class Labor Market Information System for Indonesia." World Bank Group. https://doi.org/10.1596/35378

Yildirmaz, A. 2016. "Using Big Data to Decode Private Sector Wage Growth."

Zhang, D., S. Mishra, E. Brynjolfsson et al. 2021. "The AI Index 2021 Annual Report." Stanford, CA: Stanford University. https://aiindex.stanford.edu/wp-content/uploads/2021/03/2021-AI-Index-Report_Master.pdf.

Zhu, T.J., J. Fritzler, and J. Orlowski. 2018. "Data Insights: Jobs, Skills and Migration Trends Methodology & Validation Results." World Bank Group and LinkedIn.

Zolas, N. et al. 2020. "Advanced Technologies Adoption and Use by US Firms: Evidence from the Annual Business Survey." Working Paper 28290, National Bureau of Economic Research. https://doi.org/10.3386/w28290.