

¿Qué suma la ciencia de datos a la identificación y anticipación de la demanda de habilidades?

Lucía Echeverría
Graciana Rucci

División de Mercados Laborales

NOTA TÉCNICA N°
IDB-TN-2591

¿Qué suma la ciencia de datos a la identificación y anticipación de la demanda de habilidades?

Lucía Echeverría
Graciana Rucci

Noviembre 2022

Catalogación en la fuente proporcionada por la
Biblioteca Felipe Herrera del Banco Interamericano de Desarrollo
Echeverría, Lucía.

¿Qué suma la ciencia de datos a la identificación y anticipación de la demanda de habilidades? / Lucía Echeverría, Graciana Rucci.

p. cm. -- (Nota técnica del BID ; 2591)

Incluye referencias bibliográficas.

1. Labor market-Technological Innovations-Latin America. 2. Labor market-Technological innovations-Caribbean Area. 3. Life skills-Latin America. 4. Life skills-Caribbean Area. 5. Human capital-Latin America. 6. Human capital-Caribbean Area. 7. Big data-Social aspects-Latin America. 8. Big data-Social aspects-Caribbean Area. I. Rucci, Graciana. II. Banco Interamericano de Desarrollo. División de Mercados Laborales. III. Título. IV. Serie.
IDB-TN-2591

Palabras clave: demanda de habilidades, información del mercado laboral, ciencia de datos, proyecciones de empleo.

Códigos JEL: J08, J23, J24

<http://www.iadb.org>

Copyright © 2022 Banco Interamericano de Desarrollo. Esta obra se encuentra sujeta a una licencia Creative Commons IGO 3.0 Reconocimiento-NoComercial-SinObrasDerivadas (CC-IGO 3.0 BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) y puede ser reproducida para cualquier uso no-comercial otorgando el reconocimiento respectivo al BID. No se permiten obras derivadas.

Cualquier disputa relacionada con el uso de las obras del BID que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la CNUDMI (UNCITRAL). El uso del nombre del BID para cualquier fin distinto al reconocimiento respectivo y el uso del logotipo del BID, no están autorizados por esta licencia CC-IGO y requieren de un acuerdo de licencia adicional.

Note que el enlace URL incluye términos y condiciones adicionales de esta licencia.

Las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del Banco Interamericano de Desarrollo, de su Directorio Ejecutivo ni de los países que representa.





¿QUÉ SUMA LA CIENCIA DE
DATOS A LA IDENTIFICACIÓN Y
ANTICIPACIÓN DE LA DEMANDA
DE HABILIDADES?

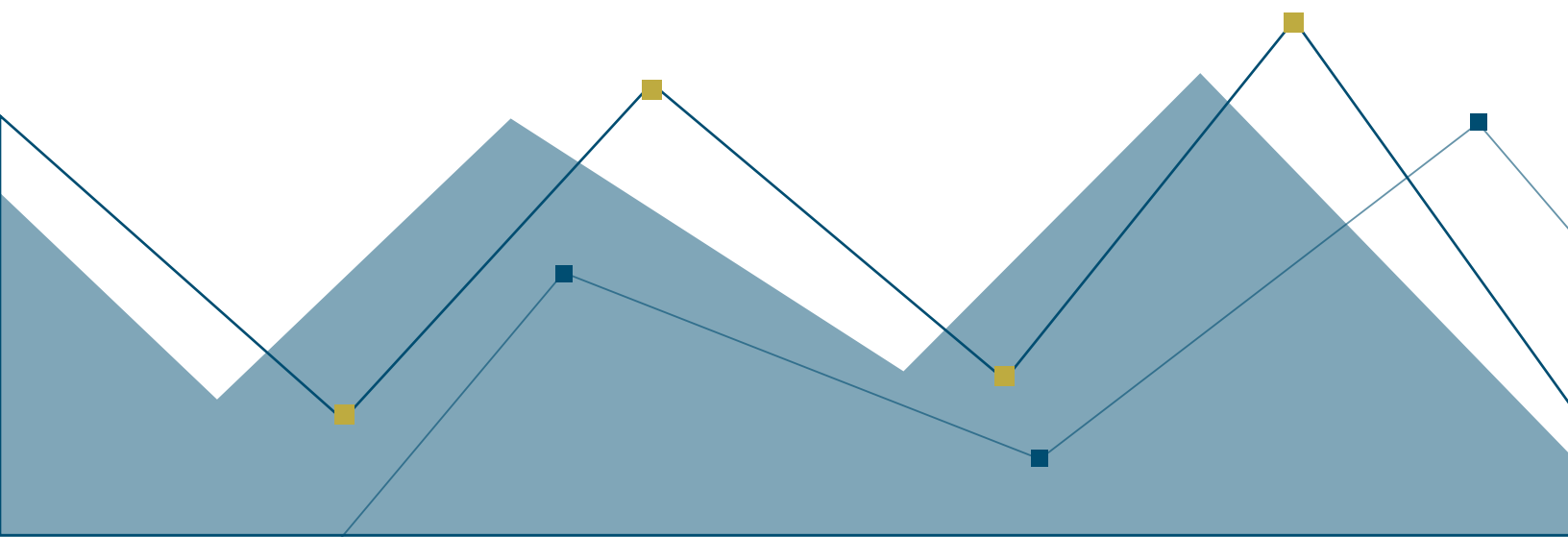
¿QUÉ SUMA LA CIENCIA DE DATOS A LA IDENTIFICACIÓN Y ANTICIPACIÓN DE LA DEMANDA DE HABILIDADES?

Lucía Echeverría y Graciana Rucci*

Octubre 31, 2022

2
< >

*Se agradecen los valiosos comentarios y sugerencias recibidas de Catterina Colombo, María Fernanda Prada y David Kaplan.



ÍNDICE DE CONTENIDOS

Índice de contenidos	3
Resumen ejecutivo	4
1. Introducción	6
2. ¿Qué son las habilidades y por qué identificar y anticipar su demanda?	8
3. Análisis basado en la ciencia de datos	11
3.1. Aspectos metodológicos: ¿Cómo se recolectan y procesan los datos masivos de vacantes online?	13
3.2. Ejemplos de países desarrollados	16
3.3. Ejemplos de América Latina y el Caribe	19
4. ¿Cómo complementar los métodos tradicionales con el análisis basado en la ciencia de datos?	22
5. Conclusiones	26
Referencias	28

RESUMEN EJECUTIVO

La identificación y anticipación de la demanda de habilidades consiste en la implementación de herramientas para producir información que permita determinar la naturaleza de las habilidades que el mercado laboral requiere, y señalar cambios en la demanda de dichas habilidades para contribuir con una toma de decisiones más informada por parte de los diferentes agentes de la economía y, así mejorar el funcionamiento del mercado laboral, entre otros.

La importancia de identificación y anticipación de la demanda de habilidades reside en que produce información para guiar las decisiones de los distintos actores de la economía. Por un lado, contribuye a que el gobierno diseñe y monitoree políticas de empleo, desarrolle una oferta de formación que se adecue a las necesidades de la demanda y asigne fondos públicos hacia ciertas certificaciones. Por otro lado, facilita que los individuos tomen decisiones informadas y estratégicas sobre su inversión en formación y trayectoria laboral, y permite que la oferta de formación se adecue a las necesidades del sector productivo.

Los métodos tradicionales se basan en análisis cuantitativos y/o en investigaciones cualitativas.¹ Los países desarrollados generan sistemas de información laboral a partir de la combinación sistemática y frecuente de distintos instrumentos que les permite contar con datos actualizados e integrados. En contraste, en los países de América Latina y el Caribe (ALC), los esfuerzos son más incipientes e incompletos.

Sin embargo, en los últimos años, el avance de la digitalización y tecnología ha promovido una nueva herramienta de gran potencial informativo y que supera muchas de las limitaciones de los métodos tradicionales: el análisis de datos masivos de vacantes de empleo *online* a través de la ciencia de datos. La gran ventaja de estos datos reside en la posibilidad de realizar un seguimiento granular, actualizado, instantáneo y con bajo costo de las tendencias que existen y emergen en la demanda de habilidades.

Al ser información masiva y no estructurada, el uso de estos datos requiere la implementación de herramientas, basadas en la ciencia de datos, que extraigan la información contenida en los sitios web de forma eficiente y automática, y la

1. Ver González-Velosa y Rucci (2016) para una descripción de los principales métodos tradicionales de análisis de demanda de habilidades.

convierta en formatos estructurados y fáciles interpretar y analizar. En muchos países desarrollados, como los de la Eurozona, Australia y Estados Unidos, estos análisis se han comenzado a implementar con frecuencia. No obstante, en América Latina aún son pocas las aplicaciones de análisis masivo de datos de demanda en el mercado de trabajo.

Esta nota técnica describe la potencialidad de la ciencia de datos para realizar análisis de identificación y anticipación de habilidades, así como su complementariedad con los métodos tradicionales. Además, en la nota se comenta algunas iniciativas recientes que pueden ser empleadas como referencia práctica y se destacan los aspectos metodológicos más relevantes que subyacen a su aplicación.

1

INTRODUCCIÓN

El creciente proceso de digitalización y los avances tecnológicos han aumentado las probabilidades de que las personas busquen trabajo a través de internet, mientras que cada vez son más los empleadores que usan las plataformas *online* para publicar las vacantes disponibles y ampliar, a un bajo costo, el alcance de sus búsquedas hacia otras regiones. Como consecuencia ha surgido una fuente de información rica, novedosa y de fácil acceso sobre el mercado de trabajo que ofrece datos masivos (*big data*) sobre las habilidades demandadas: los portales abiertos de vacantes de empleo *online*. La gran ventaja de estos datos radica en la posibilidad de contar con información que se caracteriza por ser granular, actualizada, instantánea, de alta frecuencia y recolectada a bajo costo. Esto permite realizar análisis más refinados sobre la demanda de habilidades. Dado que los datos masivos son no estructurados, su uso requiere la implementación de herramientas basadas en la ciencia de datos para extraer la información contenida en los sitios web de forma eficiente, automática y convertirla en formatos estructurados y de fácil interpretación.

El análisis de la demanda basado en los datos masivos extraídos de los portales de vacantes *online* es de especial relevancia en los países en desarrollo, como los de América Latina y el Caribe (ALC), donde existen mayores barreras, en cuanto a costos y dificultades técnicas, para implementar los métodos tradicionales de anticipación e identificación de habilidades basados en técnicas cuantitativas o cualitativas, como proyecciones, encuestas, estudios prospectivos o diálogos sectoriales. En este contexto, esta nueva fuente de datos constituye una gran oportunidad para mejorar el entendimiento sobre las dinámicas de la demanda de habilidades, dada la importancia que esto supone para el diseño de las políticas públicas. En este sentido, la información contenida en éstos análisis permite que el gobierno diseñe y monitoree políticas de empleo, y contribuye a desa-

rollar una oferta de formación que se adecue a las necesidades de la demanda, guiando las decisiones de asignación de fondos públicos e implementando instrumentos financieros que promuevan ciertas certificaciones o formaciones.

En este sentido, las incipientes iniciativas de análisis de demanda de habilidades mediante datos masivos para la región muestran el gran potencial informativo de esta herramienta y abren la puerta para su adopción y desarrollo sistemático por parte de los gobiernos.

Esta nota técnica describe la potencialidad de la ciencia de datos para realizar análisis de identificación y anticipación de habilidades, así como el grado de complementariedad con los métodos tradicionales. Asimismo, la nota comenta algunas de las iniciativas recientes que pueden ser utilizadas como referencia práctica y evalúa los principales aspectos metodológicos que subyacen a su aplicación.

La nota se estructura de la siguiente forma. En la segunda sección, se conceptualizan las habilidades y se explica la importancia de identificarlas y anticiparlas de forma sistemática y estructurada para disminuir la magnitud de los descalces en el mercado de trabajo. En la tercera sección, se explica en qué consiste el análisis de la ciencia de datos y se describen sus principales ventajas y desventajas. Asimismo, se detallan las principales cuestiones metodológicas asociadas a las experiencias recientes que han usado este análisis, tanto en países desarrollados como en América Latina y el Caribe. En la cuarta sección, se analiza en qué aspectos podría complementarse con los métodos tradicionales. Por último, se presentan las principales conclusiones.

2

¿QUÉ SON LAS HABILIDADES Y POR QUÉ IDENTIFICAR Y ANTICIPAR SU DEMANDA?

Las habilidades se definen como las capacidades, competencias, cualidades, talentos y conocimientos, ya sean innatos o adquiridos, que les permiten a los individuos contribuir productivamente con la sociedad, mejorar su bienestar y ser buenos ciudadanos (Busso et al., 2017; Prada y Rucci, 2016). La adquisición de habilidades tiene un carácter acumulativo y dinámico, ya que se desarrollan a lo largo del ciclo de vida y en distintos lugares, como el hogar, la escuela, la universidad o institutos formativos, en lugares de trabajo, entre otros. Las habilidades se pueden agrupar en tres categorías (BID, 2020):

- i. **habilidades cognitivas:** son las habilidades, como la memoria, las funciones ejecutivas y el pensamiento crítico, que se emplean para pensar, leer, aprender, recordar, razonar y prestar atención. Este concepto puede extenderse de modo de incluir las habilidades académicas (conocimiento de conceptos, procedimientos y estrategias) y digitales. Al mismo tiempo, las habilidades cognitivas pueden subdividirse en básicas, intermedias o avanzadas según su nivel de complejidad.
- ii. **habilidades socioemocionales:** son las habilidades asociadas a las cualidades individuales de las personas, incluidas sus actitudes, creencias, características de personalidad y comportamiento. También son conocidas como habilidades blandas, de comportamiento o sociales.

iii. **habilidades técnicas:** son las habilidades específicas, vinculadas con el trabajo o la ocupación de las personas. Por ello, su definición depende de la caracterización del trabajo u ocupación, de las competencias requeridas y de las tareas que involucre un determinado puesto laboral.

La identificación de la demanda de habilidades consiste en detectar los requerimientos actuales, o en el muy corto plazo, del mercado laboral. Por su parte, la anticipación de habilidades consiste en detectar las necesidades futuras, es decir, las tendencias en la demanda de habilidades que exhiben una alta probabilidad de mantenerse, emerger o desaparecer. La anticipación e identificación de habilidades demandadas es crucial en un contexto en el cual las transformaciones tecnológicas, la creciente globalización, el cambio climático y demográfico, impactan continuamente sobre las economías y los procesos productivos, ya que modifican los conocimientos y las capacidades que requieren las ocupaciones y tareas. Al mismo tiempo, estas transformaciones generan incertidumbre acerca de cuáles serán habilidades asociadas a los futuros puestos de trabajo. Tanto la oferta como la demanda de trabajo pueden adaptarse y ajustarse a estos cambios, pero lo hacen a distintos ritmos, lo que provoca un descalce de habilidades en el mercado de trabajo.² Por ejemplo, la incorporación de innovaciones en el proceso de producción de las empresas puede ser más rápido que el tiempo que requiere introducir cambios en los programas formativos para que los individuos aprendan o perfeccionen las nuevas habilidades requeridas. Esto puede generar que las personas obtengan titulaciones que no se adecuan a las nuevas capacidades que demandan los empleadores. Las asimetrías de información sobre los potenciales retornos económicos de cada ocupación, la movilidad geográfica limitada y las barreras a la participación y provisión de trayectos formativos contribuyen, también, al descalce de habilidades.

Si la magnitud del descalce entre la oferta y la demanda de habilidades es significativa y persistente puede generar costos económicos y sociales, pues la economía no estaría utilizando eficientemente su capital humano. En este sentido, el desbalance de habilidades tiene un impacto negativo en la satisfacción laboral y en los salarios de los individuos (Montt, 2015; OECD, 2014; Comyn y Strietska-Ilina, 2019), y genera menores niveles de empleo y productividad y, junto con ello, menor crecimiento económico (Sattinger, 1993; Adalet McGowan y Andrews, 2015; Comyn y Strietska-Ilina, 2019). Como consecuencia, entender las causas de dichos desbalances resulta crucial para las políticas públicas de los países, así como para el desarrollo económico y el bienestar de la población. Y, justamente

2. Para mayor detalle véanse Busso et al., 2012; Cappelli, 2014; McGuinness et al., 2018; Vandeplas y Thum-Thysen, 2019.

por esto, es fundamental que los países sigan desarrollando e invirtiendo en los métodos e instrumentos para identificar y anticipar las habilidades demandadas.

En este contexto, el desafío de identificar y anticipar tendencias en las demandas de habilidades reside en la implementación de instrumentos que conformen y alimenten sistemas de información útiles, oportunos, frecuentes y de calidad, que contribuyan a cerrar la brecha entre las habilidades demandadas y ofrecidas en el mercado laboral. Por ello, los métodos existentes buscan conocer en qué sectores y ocupaciones habrá mayor generación de empleo y cuáles son las habilidades que tendrán mayor demanda.

Cabe señalar que, en los análisis de demanda, se examinan grupos agregados de habilidades, dadas las dificultades de capturar detalladamente cada habilidad en particular. Al mismo tiempo, debido a la gran cantidad de nombres que puede recibir una misma habilidad, las clasificaciones son de vital importancia. Por otro lado, en muchas ocasiones sólo es posible anticipar la demanda de ocupaciones o cualificaciones, que suelen utilizarse como aproximaciones -proxy- de la demanda de habilidades.

La información que contienen estos instrumentos es de gran importancia para las decisiones que deben tomar los distintos actores de la economía. En este sentido, informar sobre las necesidades de los empleadores ayuda a que:

- el gobierno diseñe y monitoree políticas de empleo, como los servicios de orientación vocacional e intermediación laboral (Mazza, 2016), para: i) desarrollar una oferta de formación que se adecue a las necesidades de la demanda; ii) asignar fondos públicos hacia ciertas formaciones o certificaciones; y iii) implementar instrumentos financieros que promuevan la formación (por ejemplo, subsidios para educarse y créditos baratos para certificaciones, entre otros);
- los individuos tomen decisiones informadas y estratégicas sobre su inversión en formación y en su trayectoria laboral; y
- la oferta de formación sea pertinente, de calidad, pueda actualizarse según las necesidades de habilidades del sector productivo (Amaral et al., 2017).

3

ANÁLISIS BASADO EN LA CIENCIA DE DATOS

La principal fuente de información masiva que permite realizar análisis de demanda consiste en los datos relativos a las vacantes publicadas en las plataformas de redes profesionales y los portales *online*.³ La información contenida en los anuncios de los portales incluye datos estructurados sobre las características de la vacante, por ejemplo, el tipo de contrato, los horarios, el rango de salario, su ubicación geográfica y la categoría ocupacional o el sector al que pertenece el puesto de trabajo. También puede incluir campos de texto no estructurados que contienen las habilidades específicas y el nivel de experiencia requeridos para el puesto de trabajo, así como las titulaciones o el nivel de calificación deseado. También se puede describir la empresa y las características de la modalidad de trabajo (OIT, 2020; Cedefop, 2021). El Diagrama 1 sintetiza el tipo de información que contienen los anuncios sobre las vacantes *online*.

DIAGRAMA 1. INFORMACIÓN DISPONIBLE EN LAS VACANTES DE TRABAJO ONLINE

DATOS ESTRUCTURADOS	DATOS NO ESTRUCTURADOS
Tipo de contrato	Habilidades específicas requeridas
Horarios	Nivel de experiencia requerido
Salario	Titulaciones o el nivel de calificación
Ubicación geográfica	Características de la empresa
Categoría ocupacional o sector	Características de la modalidad de trabajo

Fuente: Elaboración propia en base a OIT (2020), Cedefop (2021) y BID (2022).

3. Si bien el foco principal del análisis de datos masivos para identificar y anticipar demandas de habilidades se centra en los portales de vacantes *online*, también existen otras fuentes de información disponibles que podrían emplearse como las patentes, las bases de datos científicas o la publicación de cursos *online*. Para más detalles sobre el empleo de estos datos ver Cedefop (2021).

Ventajas

El empleo de estos datos posee importantes ventajas. En primer lugar, dado que la información ya se encuentra *online*, su uso es eficiente en términos de volumen de datos y costos. Si bien el desarrollo de los algoritmos iniciales no es instantáneo, una vez establecidos, pueden ejecutarse de forma rápida y sin costos adicionales, generando información que puede ser usada de forma casi inmediata. En segundo lugar, permiten análisis más eficientes y comprehensivos debido a que –a partir de las variables recolectadas– entregan información granular que facilita el mapeo de las habilidades.

En tercer lugar, como la información se actualiza de forma frecuente, es posible identificar tendencias, analizar cambios en la demanda de empleo, ocupaciones y habilidades, así como detectar qué habilidades tienen riesgo de quedar obsoletas, cuáles son transversales o comunes a más de una ocupación, o bien anticipar aquellas que se encuentran en expansión. En esta línea, el uso de datos masivos permite comprender cómo varían los requerimientos de habilidades asociados a los puestos de trabajo. De este modo se genera un “GPS del empleo” que ayuda a los trabajadores a identificar cuáles son las habilidades que le permitirán pasar de una ocupación en declive a otra emergente (Amaral et al., 2018; OECD, 2018). Esto es especialmente pertinente en contextos de incertidumbre sobre el futuro, especialmente en relación a las innovaciones tecnológicas o la evolución de la demanda.

Limitaciones

Este tipo de análisis también posee un conjunto de limitaciones. En primer lugar, las habilidades que se incluyen en el aviso de una vacante pueden no reflejar completamente el perfil deseado, puesto que solo suelen incluirse aquellas habilidades o calificaciones más relevantes para filtrar las postulaciones. También es frecuente que las empresas publiciten el mismo puesto en distintos portales, lo cual genera la necesidad de contar con algoritmos que puedan depurar correctamente la información duplicada. En segundo lugar, pueden surgir dificultades para efectuar comparaciones internacionales, dado que los datos pueden ser recolectados en distintos portales y en distintos idiomas, lo cual requiere un esfuerzo adicional en la estructuración de la información. En este sentido, pueden existir sesgos involuntarios en la agrupación de los datos. Por ejemplo, para que los algoritmos de los programas puedan leer y coleccionar la información contenida en las vacantes, será necesario, en muchos casos, emplear un vocabulario estandarizado, lo cual puede incluir decisiones para simplificar u homogeneizar dicha información (Di Capua et al., 2020).

En tercer lugar, los datos no suelen ser representativos de todas las vacantes del mercado de trabajo. En efecto, es posible que se publiciten más aquellas vacantes que requieren perfiles de trabajadores relativamente más calificados y que se genere una sobrerrepresentación de determinados sectores u ocupaciones. Similarmente, los anuncios suelen concentrarse en puestos de trabajo pertenecientes a la economía formal, por lo tanto, las habilidades requeridas en el mercado informal se ven subrepresentadas. De igual modo, la representatividad regional puede variar, ya que el uso de los portales puede ser heterogéneo aún dentro de un mismo país. En cuarto lugar, existen potenciales problemas con respecto a las políticas de privacidad y uso de los datos (Kureková et al., 2014; OIT, 2020; Cedefop, 2021).

En síntesis, las nuevas tecnologías y la ciencia de datos contribuyen con el análisis de demanda de habilidades y lo harán cada vez más en un contexto de crecientes avances tecnológicos. Sin embargo, su implementación requiere contar con la supervisión humana, y de herramientas estadísticas, para establecer criterios de recolección, comparación, homologación y clasificación de los datos masivos.

3.1. Aspectos metodológicos: ¿Cómo se recolectan y procesan los datos masivos de vacantes online?

El sistema de producción de los datos masivos que provienen de las vacantes de trabajo contenidas en las plataformas *online* se lleva a cabo en distintas fases (OIT, 2020; Cedefop, 2021), resumidas en el Diagrama 2. Las mismas consisten en:

- i. Minería de datos de las vacantes:** consiste en la selección de los portales o las plataformas para identificar las vacantes disponibles en internet y descargar su contenido. Se pueden emplear distintas técnicas para tal efecto:
 - a.** Web scraping se emplea para recolectar datos de las páginas web que ya se encuentren estructurados, a fin de extraerlos conociendo previamente la posición exacta de la información. Por este motivo, es posible que deban programarse web scrapers específicos para cada portal.
 - b.** Web crawling emplea un robot programado para buscar sistemáticamente información no estructurada en las plataformas *online* y descargarla. Este método es más general y fácil de desarrollar que el web scraping. Sin embargo, también puede recolectar datos con contenido irrelevante que generen ruido y requieran una mayor depuración en la siguiente fase.

c. Acceso directo vía la interfaz de programación: permite descargar la información referida a las vacantes directamente desde el portal. Los datos recolectados mediante esta vía son de mayor calidad que los obtenidos mediante el web scraping, y pueden ser descargados con más rapidez. Sin embargo, dado que este método implica el acceso directo a los datos, requiere acuerdos y la autorización expresa por parte de la plataforma y, en muchos casos, puede conllevar costos adicionales.

ii. Pre-procesamiento: consiste en la depuración de los datos recolectados e involucra los siguientes procesos:

a. Homologación del idioma, consiste en agrupar palabras/léxicos o en emplear un vocabulario estandarizado en el caso en que se descargan datos provenientes de distintos portales que utilizan distinto idioma. Esto es particularmente importante cuando el análisis incluye a varios países y/o busca realizar comparaciones internacionales.

b. Limpieza de datos, para eliminar la información que no sea relevante para el análisis de las habilidades, como las publicidades, elementos de presentación, logos e imágenes.

c. Combinación de los datos provenientes de distintos portales, dado que una empresa puede publicitar la misma vacante en más de una plataforma. Si bien la duplicación no es deseada, en una primera instancia puede aumentar la información recolectada ya que pueden existir ciertos datos asociados a un mismo puesto de trabajo que pueden ser publicitados en un sitio, pero no en otro, permitiendo complementar e incrementar el detalle de la información.

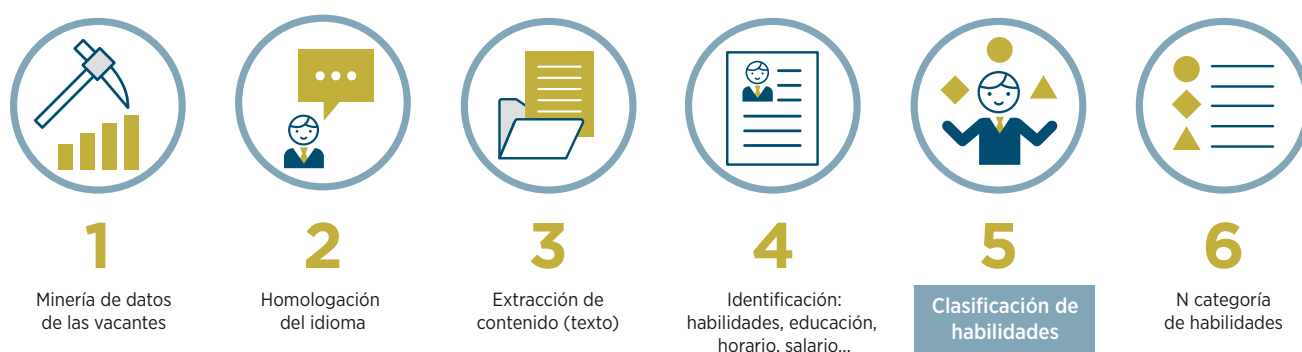
d. Eliminación de datos sobre vacantes duplicadas, para lo cual es necesario establecer criterios que permitan identificar esas vacantes (por ejemplo, la ubicación y las características de la vacante).

iii. Extracción de contenido: consiste en la extracción de los datos de interés mediante algoritmos de machine learning que relacionan el contenido de los portales con una determinada ontología con el propósito de ordenar los datos.

iv. Identificación del contenido: consiste en la identificación de los distintos tipos de información contenida en las vacantes, por ejemplo, información sobre las habilidades requeridas, el nivel educativo, la experiencia, las características del puesto de trabajo, entre otras (ver Diagrama 1).

- v. Clasificación de las habilidades:** consiste en la organización de la información identificada ya que la misma se encuentra no estructurada. En esta etapa, es fundamental el proceso de selección e implementación de taxonomías que permitan agrupar la información granular recolectada, tal de estructurarla en grupos de habilidades y ocupaciones, permitiendo su posterior análisis (Altamirano et al., 2019; Altamirano y Amaral, 2020).
- vi.** Como resultado final, es posible obtener una base de datos que contiene información armonizada sobre distintas características de las vacantes *online*, incluyendo las categorías de habilidades previamente definidas.

DIAGRAMA 2. RESUMEN DE LA METODOLOGÍA PARA RECOLECTAR Y PROCESAR DATOS MASIVOS



Fuente: BID (2022)

Cabe destacar que el avance de la tecnología y la ciencia de datos ha permitido, también, la creación de sitios web, como es el caso de *Google Jobs*⁴, que recopilan y agregan los datos de vacantes abiertas publicadas en distintas partes del mundo, a la vez que, depuran, limpian y agrupan la información para que pueda ser utilizada directamente por los usuarios, ahorrando los procesos ii.c y ii.d descritos anteriormente.

4. Ver <https://www.cedefop.europa.eu/en/tools/skills-online-vacancies>.

3.2. Ejemplos de países desarrollados

A continuación, seleccionamos y comentamos algunas iniciativas recientes de análisis, realizados tanto por organismos gubernamentales como por entidades privadas, de demanda de habilidades basados en la ciencia de datos que pueden ser utilizadas como referencia práctica, destacando los principales aspectos metodológicos, tales como la elección de los portales, la extracción de la información, la estructuración de los datos y las taxonomías empleadas para los análisis de la demanda de habilidades.⁵

“Skills-OVATE”- Cedefop

En Europa, la iniciativa pionera del análisis de la demanda de habilidades a partir de información de vacantes extraída de los anuncios *online* fue lanzada en 2021 por el Centro Europeo para el Desarrollo de la Formación Profesional (Cedefop), un organismo de la Unión Europea. Esta herramienta, denominada “Skills-OVATE” (“Skills Online Vacancy Analysis Tool for Europe”)⁶ ofrece una base de datos detallada, granular y casi en tiempo real sobre los puestos de trabajo y las habilidades demandadas en 28 países europeos desde 2018.

Los datos son extraídos de millones de vacantes publicitadas en distintas fuentes, como portales de empleo privados y públicos, agencias de reclutamiento, diarios *online* y páginas web corporativas. Los datos sobre el mercado de trabajo y las tendencias de las habilidades se actualizan cuatro veces por año. Skills-OVATE utiliza una combinación de tres métodos para recolectar información sobre las vacantes: acceso directo vía la interfaz de programación, web scraping y web crawling. En este sentido, como parte de la metodología empleada por Cedefop, se informa a los propietarios de los portales de vacantes acerca de la extracción de datos prevista y, en caso de ser posible, se celebran acuerdos formales para enmarcar la cooperación entre las partes. Esto posibilita que en muchos casos los propietarios otorguen acceso directo a los datos a través de interfaz de programación de la página (Cedefop, 2019). Una vez extraída, la información sobre las vacantes se clasifica de acuerdo con taxonomías internacionales de las ocupaciones y habilidades. En particular:

5. Para más experiencias sobre el uso de datos masivos para analizar la demanda de habilidades en países desarrollados se puede consultar, por ejemplo, la guía elaborada por la European Training Foundation (Mezzanzanica y Mercurio, 2019).

6. Ver <https://www.cedefop.europa.eu/en/tools/skills-online-vacancies>.

i. para las ocupaciones se emplea la International Standard Classification of Occupations (ISCO-08, CIUO-08 según sus siglas en español), elaborada por la Organización Internacional del Trabajo (OIT) que divide los empleos en función del tipo de trabajo realizado y de las competencias requeridas para cumplir satisfactoriamente con las obligaciones del puesto.

ii. para las habilidades se emplean dos clasificaciones alternativas. Por un lado, se utiliza la European Classification of Occupations and Skills (ESCO versión 1) que clasifica habilidades, competencias, cualificaciones y ocupaciones y provee una taxonomía común para los países de Europa. Por otro lado, se utiliza la Occupational Information Network (O*NET), elaborada por el Departamento de Empleo de EE.UU., que clasifica cientos de ocupaciones en función de cuatro dimensiones: a) las habilidades y los conocimientos que se necesitan para realizar un trabajo; b) las características personales y los intereses y valores que requiere una ocupación; c) los años de experiencia, entrenamiento y capacitación relativos al puesto; y d) los requerimientos de tareas y actividades del puesto así como los contextos donde se desempeña el trabajo (factores físicos, sociales y organizacionales).

“Job Vacancy Index”- Australia

También se destaca la experiencia en Australia, donde la Comisión Nacional de Habilidades del gobierno elabora un reporte mensual, denominado “Job Vacancy Index” sobre la base del análisis de vacantes publicadas en tres plataformas, “SEEK”, “Career One” y “Australian Job Search”.⁷ Los resultados se presentan como una serie temporal para estudiar la evolución de vacantes diferenciadas por región geográfica, estados, ocupación y habilidades. La información sobre las vacantes es codificada de la siguiente forma:

i. las ocupaciones se clasifican de acuerdo con la taxonomía desarrollada por Australia y Nueva Zelanda, “Australian and New Zealand Standard Classification of Occupations” (ANZSCO).

ii. las habilidades asociadas a los puestos de trabajo se clasifican según el nivel educativo o la experiencia normalmente requerida para trabajar en la ocupación.

7. Metodología disponible en: <https://labourmarketinsights.gov.au/our-research/page/internet-vacancy-index-methodology>

Burning Glass Technologies (BGT)

Por su parte, la empresa privada Burning Glass Technologies (BGT) analiza miles de vacantes en distintos países para identificar cuáles son los empleos más demandados, cuáles son las habilidades que demandan los empleadores y cuáles son las carreras que ofrecen el mayor potencial para los trabajadores. Por ejemplo, en Cammeraat y Squicciarini (2021) se presenta un análisis realizado por BGT para seis economías, Australia, Canadá, Nueva Zelanda, Singapur, Reino Unido y Estados Unidos, con datos del período 2010-2019. En los análisis de BGT, la extracción de datos sobre puestos de trabajo publicitados *online* se realiza mediante la técnica de web scraping diario en 40.000 bolsas de trabajo y sitios web de empresas. BGT efectúa una depuración de las vacantes duplicadas y analiza el texto de los anuncios para categorizar las vacantes *online* y estructurarlas en función de su ubicación geográfica, ocupación, industria, habilidades y educación requerido, y niveles de experiencia. Las ocupaciones y habilidades se codifican de acuerdo con distintas taxonomías:

- i. las ocupaciones se agrupan de acuerdo a la clasificación propia de BGT denominada “BGTOcc”, que está disponible para todos los países. Por otro lado, se emplean las clasificaciones internacionales y nacionales empleadas comúnmente, tal como O*NET, Standard Occupational Classification (SOC) y ANZSCO.
- ii. las habilidades se agrupan de acuerdo a la clasificación propia de BGT que consiste en tres categorías (BGT, 2019): habilidades especializadas o técnicas (por ejemplo, soldadura, desarrollo de software y análisis financiero), habilidades básicas (por ejemplo, comunicación, solución de problemas y creatividad) y habilidades de software (por ejemplo, Adobe Photoshop and AutoCAD).

3.3. Ejemplos de América Latina y el Caribe

Si bien hay estudios y análisis recientes para ALC, los sistemas de información laboral públicos (o de la política pública) con capacidad para aprovechar la potencialidad de datos masivos de vacantes *online* y usar ciencia de datos para la toma de decisiones, están en construcción. Esto contrasta con la experiencia de los países desarrollados, donde los análisis de habilidades son realizados de forma sistemática por los gobiernos estatales y, en menor medida, por entidades privadas.

El Cuadro 1 describe las características principales de algunos estudios sobre demanda de trabajo con datos masivos que se han realizado recientemente en países de América Latina. En particular, se presentan los métodos de extracción de datos empleados, aspectos relevantes sobre las vacantes analizadas, la taxonomía de ocupaciones y/o habilidades empleadas, así como las principales lecciones aprendidas. En relación a esto último, estos estudios permiten ilustrar las ventajas de utilizar datos de vacantes *online*. Por ejemplo, se observa que estos trabajos han procesado y analizado un gran volumen de información, lo cual únicamente es posible mediante el uso de la ciencia de datos. Asimismo, los estudios destacan la importancia de estos análisis para realizar seguimiento de tendencias, ya que es posible actualizar la información frecuentemente, por ejemplo, mensualmente. Gracias a ello, varios estudios han permitido elaborar portales interactivos que visualizan la información generada y actualizada sobre la demanda de trabajo, a los cuales pueden acceder tanto trabajadores como hacedores de política. También se destaca la importancia de estos análisis para identificar cuáles son las ocupaciones emergentes y los cambios de composición de habilidades dentro de una misma ocupación. Asimismo, en el caso de emplear aquellos portales que disponen tanto de información de demanda como de oferta, es posible generar una única base de datos que permita caracterizar el mercado.

Por otro lado, estos estudios nos permiten comprender mejor cuáles son los sectores de trabajadores que son mejor representados al utilizar datos de vacantes *online*, por ejemplo, trabajos de alta calificación o sectores de servicios o profesionales y técnicos. Por último, se destaca la necesidad de que los organismos públicos posean las herramientas legales que les faciliten una mejor descarga de información de los portales *online*.

CUADRO 1. ESTUDIOS SOBRE DEMANDA DE TRABAJO CON DATOS MASIVOS EN AMÉRICA LATINA Y EL CARIBE

ESTUDIO/ ORGANISMO	PAÍSES	MÉTODO DE EXTRACCIÓN DE DATOS	ASPECTOS SOBRE LAS VACANTES ANALIZADAS	CLASIFICACIÓN DE OCUPACIONES/ HABILIDADES	LECCIONES APRENDIDAS
Banco Central de Chile (Arraño y Jara, 2019)	Chile	• <i>web scraping</i> de 3 portales de vacantes	<ul style="list-style-type: none"> • se consideran 3 portales que agrupan el 90% de los avisos • 2 concentran la demanda de alta calificación 	i) demanda de trabajo según cargo y región, ii) ocupaciones clasificadas según CIUO-88	<ul style="list-style-type: none"> • se publica mensualmente un “Índice de Avisos Laborales de Internet” que proporciona una medición oportuna de la demanda de trabajo
Amaral <i>et al.</i> (2018) BID	Argentina, Brasil, Chile y México	• <i>web scraping</i> de la red profesional LinkedIn	<ul style="list-style-type: none"> • la información contenida en LinkedIn representativa de los segmentos más calificados 	i) se construye una taxonomía con la información de las tareas que especifican los usuarios en sus perfiles	<ul style="list-style-type: none"> • se destaca el uso de estos análisis como “GPS del empleo” para detectar las habilidades que permitirían pasar de una ocupación en declive a otra emergente
González-Velosa y Peña (2019) BID y LACEA	Argenti- na, Chile, Colombia, México y Perú	• <i>web scraping</i> de 12 portales de vacantes	<ul style="list-style-type: none"> • 6,9 millones de vacantes 	i) ocupaciones y habili- dades según O*NET y el marco conceptual de Autor <i>et al.</i> (2013) ^a	<ul style="list-style-type: none"> • se destaca la posibilidad de caracterizar empleos que constituyen un porcentaje muy importante del total de empleos en cada país (sector servicios y comercio, con alta rotación y baja remuneración)
Di Capua <i>et al.</i> (2020) BID	Uruguay	• <i>web scraping</i> de 3 portales de vacantes	<ul style="list-style-type: none"> • 145.969 vacantes 	i) ocupaciones según CIUO-08, ii) habilidades de acuerdo con O*NET y el marco conceptual de Acemog- lu y Autor (2011) ^b , iii) experiencia, educación y entrenamiento según O*NET ^c	<ul style="list-style-type: none"> • se destaca que las empresas recurren relativamente más a portales de empleo online para contratar trabajadores con un perfil profesional y técnico

¿QUÉ SUMA LA CIENCIA DE DATOS A LA IDENTIFICACIÓN Y ANTICIPACIÓN DE LA DEMANDA DE HABILIDADES?

ESTUDIO/ ORGANISMO	PAÍSES	MÉTODO DE EXTRACCIÓN DE DATOS	ASPECTOS SOBRE LAS VACANTES ANALIZADAS	CLASIFICACIÓN DE OCUPACIONES/ HABILIDADES	LECCIONES APRENDIDAS
OIT/Cinterfor (2022)	Uruguay (2019)	• <i>web scraping</i> de la plataforma “Busco Jobs”	• se valora la consistencia temporal de “Busco Jobs” en términos de su correlación positiva con la tasa de actividad, empleo y con indicadores de la Seguridad Social	i) ocupaciones de acuerdo con la clasificación internacional CIUO-08	• se destaca como ventaja la posibilidad de vincular los datos de demanda con los de oferta laboral desde una única base de datos
	Paraguay (2021)	• <i>web scraping</i> de portales de empleo públicos y privados	• 3.185 vacantes	i) ocupaciones clasificadas según las ocupaciones críticas del Banco Mundial ^d	• se destaca la representación de los datos mediante un tablero de visualización, como un producto de fácil acceso para los individuos
	República Dominicana (2021)	• <i>web scraping</i> de 3 portales de vacantes y el portal oficial de empleo del Ministerio del Trabajo	• 11.000 vacantes	i) ocupaciones clasificadas según CIUO-08, ii) habilidades según blandas y digitales	• el ejercicio recalca la necesidad de dotar al Ministerio de Trabajo con herramientas legales que lleven a consolidar la descarga de información estructurada de las vacantes
Gontero y Menéndez (2021) Hilbert y Lu (2020) CEPAL	33 países de América Latina y el Caribe	• <i>web scraping</i> y <i>web crawling</i> (según la estructura de la información) de 8 portales locales y globales	• información disponible sobre 1.395.333 vacantes diarias • 2 de los portales disponen de información de oferta de trabajadores (<i>Freelancer y Profdir</i>)	i) ocupación clasificadas según la codificación CIUO-08	• se destaca la construcción de un portal interactivo web que contiene información semanal sobre la demanda de vacantes de empleo en cada país

Fuente: elaboración propia.

Nota: ^a esta clasificación permite agrupar las ocupaciones según los tipos de tareas: cognitivas no rutinarias analíticas; cognitivas no rutinarias interpersonales; cognitivas rutinarias, manual rutinarias y manual físicas no rutinarias. ^b Este marco conceptual clasifica las habilidades según sean cognitivas no rutinarias analíticas; cognitivas no rutinarias interpersonales; cognitivas rutinarias, manual rutinarias y manual físicas no rutinarias. ^c Los requerimientos de pueden agrupar según impliquen poca o ninguna preparación necesaria, alguna preparación necesaria, preparación media necesaria, considerable preparación necesaria y gran preparación necesaria. ^d Esta clasificación considera ocupaciones que cumplen con tres criterios (requieren habilidades, son demandadas y son estratégicas), y ha sido empleada en estudios de Indonesia y Malasia y, más recientemente, en Australia, Colombia, Estados Unidos, Reino Unido y Turquía.

4

¿CÓMO COMPLEMENTAR LOS MÉTODOS TRADICIONALES CON EL ANÁLISIS BASADO EN LA CIENCIA DE DATOS?

22

< >

Los métodos para identificar la demanda de habilidades existente y anticipar posibles cambios y tendencias, tradicionalmente se han basado en análisis cuantitativos y/o en investigaciones cualitativas.⁸

Los enfoques cuantitativos, por lo general, emplean variables -como las ocupaciones, el nivel de cualificaciones o el nivel educativo- para aproximar las habilidades. Son indicadores indirectos cuyo objetivo es establecer predicciones generales a partir de información a nivel macro. Sin embargo, este tipo de análisis no necesariamente arroja resultados lo suficientemente informativos y específicos para guiar las decisiones de política pública. Estas fuentes pueden ser muy ricas en información, pueden ser representativas a nivel nacional y posibilitar análisis de largo plazo; sin embargo, las encuestas requieren tiempo para su elaboración e implementación, así como para el procesamiento de los datos, previo a la publicación de los resultados.

Por su parte, las técnicas cualitativas permiten obtener información más detallada, ya que permiten identificar habilidades o necesidades de competencias específicas a nivel regional, sectorial o local, y para ciertas ocupaciones en particular. Sin embargo, demandan bastante tiempo de implementación y procesamiento,

8. Ver González-Velosa y Rucci (2016) para una nota más comprehensiva sobre estos métodos.

son costosas, y existe el riesgo de que la información que producen quede obsoleta antes de ser utilizada. Además, en las economías menos avanzadas, es posible que la implementación de métodos cuantitativos y cualitativos no pueda realizarse con la regularidad necesaria debido a su alto costo. En este sentido, las limitaciones de estos métodos se contraponen con las ventajas de los análisis basados en la ciencia de datos (ver Sección 3).

El Cuadro 2 sintetiza los requerimientos de datos, así como las principales ventajas y desventajas de los enfoques tradicionales (instrumentos de proyección, encuestas a empleadores, sistemas de seguimiento de vacantes, estudios que combinan técnicas cuantitativas y cualitativas, y estudios prospectivos) y de los análisis basados en ciencia de datos.

La implementación de uno u otro enfoque dependerá de múltiples aspectos: la disponibilidad de recursos monetarios y técnicos, los objetivos del análisis, la participación de los distintos sectores, la gestión institucional, entre otros. Cada método genera información distinta en términos de granularidad, horizonte temporal y frecuencia de actualización de los datos. También se pueden alcanzar distintos niveles de análisis según el alcance informativo que tenga método escogido, el cual está será determinado en parte por el propósito de política del estudio (Van Breugel, 2017).

Los instrumentos cuantitativos pueden ser de alcance nacional, regional o sectorial/ocupacional. Si bien la mayor parte de los países desarrollados realizan análisis con proyección nacional, suelen complementarlo con resultados diferenciados por estados, provincias o regiones. Por su parte, los estudios cualitativos, entregan información de carácter más desagregado o específico, ya sea regional, sectorial o local. Sin embargo, ninguno de estos instrumentos es capaz de generar por sí solo toda la información necesaria para guiar adecuadamente las decisiones individuales y las políticas públicas de formación y empleo. Por lo tanto, los métodos cuantitativos y cualitativos pueden complementarse, de modo tal que las desventajas de un enfoque se mitiguen al combinarse con otro.

En este contexto, los análisis que emplean datos masivos se presentan como una oportunidad para enriquecer los métodos tradicionales, ya que generan información de distinta naturaleza que puede complementar la de los enfoques tradicionales. Uno de sus principales atributos es la generación de información en tiempo real, de forma rápida, casi instantánea, y con bajo costo. Esto permite superar una de las principales barreras para la implementación regular de los métodos tradicionales. En contraste, la identificación y anticipación de habilidades a partir de la ciencia de datos en los países en desarrollo puede verse limitada por

el menor uso de portales de vacantes online en determinados sectores, como por ejemplo para los empleos de baja calificación; o bien por la magnitud del mercado de trabajo informal, dado que es esperable que las vacantes online representen únicamente el sector formal. En este sentido, la baja representatividad de las vacantes online se presenta como una de las mayores limitaciones del uso de los datos masivos. Por lo tanto, sería importante que la información que no puede capturarse a través de los datos masivos, podría ser complementada con aquella recolectada mediante otros enfoques.

En la actualidad, son pocas las iniciativas que combinan los métodos tradicionales con el análisis de datos masivos basado en la ciencia de datos. Esfuerzos recientes han sido implementados por la OIT en áreas y contextos de estudio específicos (OIT, 2020). En uno de ellos, se analizan las habilidades en el marco de la transición hacia economías descarbonizadas y sustentables en términos ambientales para 32 países (OIT, 2019).

Este proyecto, denominado “Skills for a greener future” (“Habilidades por un Futuro más verde”), involucró la combinación de técnicas cualitativas, cuantitativas y el análisis de datos masivos. En particular, se implementó un análisis cualitativo a partir de estudios nacionales en profundidad y de encuestas realizadas a un grupo de expertos. Este análisis tuvo por objetivo examinar las políticas relacionadas con el medioambiente y su compatibilidad con aquellas destinadas al desarrollo de habilidades y con los efectos de las actividades económicas. Asimismo, para complementar los estudios cualitativos, se estimó la demanda de habilidades bajo dos escenarios globales: sustentabilidad energética y economía circular. Dicha estimación se efectuó con un modelo insumo-producto multirregional que permitió analizar las transacciones entre 163 industrias en 44 países. Este modelo explora cuál es el impacto estimado sobre el empleo para el año 2030 en caso de que el aumento de la temperatura global continúe siendo menor a 2°C. No obstante, dado que la información sobre ocupaciones son aproximaciones imperfectas e insuficientes para medir las habilidades y competencias, se incluyó un análisis de datos masivos, elaborado por BGT, sobre las tareas y habilidades requeridas en las vacantes publicitadas online. Este análisis permitió identificar la composición de habilidades asociadas a las distintas ocupaciones con un grado de desagregación imposible de generar con los métodos tradicionales de análisis.

En resumen, uno de los principales desafíos para avanzar en la identificación y anticipación de la demanda de habilidades radica en cómo compatibilizar, de la forma más eficiente y efectiva posible, la amplia variedad de datos y técnicas disponibles, incorporando sistemáticamente el análisis de datos masivos.

CUADRO 2. RESUMEN DE INSTRUMENTOS PARA IDENTIFICAR Y ANTICIPAR DEMANDAS

INSTRUMENTO	REQUERIMIENTO DE DATOS	VENTAJAS	DESVENTAJAS
Modelos de pronósticos (cuantitativo)	<ul style="list-style-type: none"> • Información confiable y consistente a lo largo del tiempo sobre mercados de trabajo y características de la población 	<ul style="list-style-type: none"> • Integrales • Consistentes • Corto y largo plazo • Supuestos explícitos y transparentes • Reproducibilidad 	<ul style="list-style-type: none"> • Exigencia de datos • Costoso • Calidad depende de los datos y de supuestos • Pueden dar falsa impresión de precisión • Resultados pueden ser generales • No todo es cuantificable
Encuestas a empleadores (cuantitativo)	<ul style="list-style-type: none"> • Registros de la empresa para el diseño muestral • Recolección primaria de datos 	<ul style="list-style-type: none"> • Participación directa del empleador • Fáciles de administrar e implementar • Posibilidad de enfocarse en el comportamiento (en vez de en percepciones) • Mide las habilidades de forma directa 	<ul style="list-style-type: none"> • Inconsistencia entre sectores • Subjetividad en las respuestas • Corto plazo • Tasas de respuesta bajas • Costoso si se implementa en muestras grandes para obtener representatividad • Requiere conocimiento de habilidades por parte del encuestado
Sistemas de seguimiento de vacantes (cuantitativo)	<ul style="list-style-type: none"> • Datos administrativos consistentes y representativos 	<ul style="list-style-type: none"> • Participación directa del empleador • Información sectorial holística • Puede permitir abordar problemas en profundidad • Participación directa del empleador • Fáciles de administrar, interpretar y actualizar • Refiere a trabajos disponibles • Objetivo 	<ul style="list-style-type: none"> • Cobertura parcial: no es representativo de toda la demanda • Corto plazo • Posibles inconsistencias temporales
Estudios sectoriales y regionales (cuantitativo y cualitativo)	<ul style="list-style-type: none"> • Datos sectoriales • Enfoque cualitativo no requiere datos (e.g. focus groups, entrevistas) 	<ul style="list-style-type: none"> • Participación directa del empleador • Información sectorial holística • Puede permitir abordar problemas en profundidad 	<ul style="list-style-type: none"> • Análisis parcial • Predicciones inconsistentes entre sectores • No sistemáticos • Puede ser subjetivo, no representativo y no explícita supuestos
Estudios prospectivos (cualitativo)	<ul style="list-style-type: none"> • Diversos insumos como reportes, resultado de proyecciones cuantitativas, información del mercado laboral o estudios sectoriales 	<ul style="list-style-type: none"> • Participación directa del empleador • Holístico • Permite abordar problemas en profundidad • Incorpora la incertidumbre sobre el futuro 	<ul style="list-style-type: none"> • Puede ser no sistemático • Puede ser inconsistente • Puede ser subjetivo
Análisis basado en ciencia de datos	<ul style="list-style-type: none"> • Información online en tiempo real sobre vacantes 	<ul style="list-style-type: none"> • Gran volumen de información actualizada y a con bajo costo • Información granular • Seguimiento de tendencias emergentes en la demanda de ocupaciones y habilidades, y de los cambios de composición de habilidades dentro de una ocupación 	<ul style="list-style-type: none"> • Muestra puede no ser representativa de todo el mercado laboral • Potencial sesgo en las habilidades requeridas para las vacantes • Información imperfecta y no estructurada • Errores de medición (duplicación) • Problema de privacidad de los datos • Habilidades estadísticas para el análisis de datos

5

CONCLUSIONES

En esta nota se expone y debate sobre las posibilidades que ofrece la ciencia de datos para mejorar la identificación y anticipación de la demanda de habilidades. En este sentido, sobresale su potencial para producir información que complemente el desarrollo de sistemas modernos de información laboral que contribuyan a un mejor funcionamiento del mercado de trabajo.

También se describen las ventajas y desventajas de los datos masivos de portales de vacantes abiertos online, así como los aspectos metodológicos de mayor relevancia. Dado que esta técnica es relativamente nueva, es conveniente monitorear sus avances y asegurar que su uso sea ético y apropiado, a fin de continuar aprovechando el valor de los datos disponibles en las plataformas de vacantes online. Más aún, es posible que se siga avanzando en superar las limitaciones que hoy se encuentran en este tipo de análisis, como la representatividad, los posibles sesgos, el acceso y la privacidad de los datos recolectados y otras cuestiones éticas. Por ejemplo, el alcance a los sectores hoy no representados o grupos más vulnerables puede mitigarse mediante la implementación conjunta con métodos tradicionales y/o a partir de la búsqueda de nuevas fuentes de información como, por ejemplo, las redes sociales o comunidades virtuales de cámaras empresariales, sindicatos, entre otras.

Una gran ventaja existente para aprovechar más y mejor las nuevas tecnologías y la ciencia de datos para la política pública es la conjunción dada por: la presencia de los gigantes de las tecnologías, como Google, que realizan grandes inversiones y avances acelerados, los aprendizajes de países desarrollados e iniciativas privadas que están adelantados en su aplicación, y la investigación de frontera que genera la academia. Por tal motivo, los hacedores de política pública de América Latina y el Caribe se encuentran en un escenario en el que no requie-

ren crear nuevas herramientas de análisis en las instancias públicas, sino alinear y adaptar esfuerzos para implementar la tecnología que ya existe, consolidar equipos técnicos y desarrollar capacidades adecuadas en un buen uso de nuevas tecnologías y la ciencia de datos.

En este marco, clave que la política pública consolide estrategias y acuerdos para generar sistemas de información, articulando esfuerzos con el sector privado en concordancia con las características de gobernanza e idiosincrasia propias de cada país. Estas estrategias pueden involucrar que los gobiernos facilitan que las entidades privadas produzcan información sobre la demanda de habilidades, que se celebren acuerdos de cooperación con entidades privadas o que los gobiernos incorporen la generación y provisión de esta información a la propia operación del sector público.

En otras palabras, nos encontramos ante oportunidades únicas para aprovechar y apalancar esfuerzos público-privados orientados, en este caso, a los mercados laborales y la demanda de habilidades. Dichos impulsos pueden generar sinergias que permitan aprovechar la variedad de datos e instrumentos para producir sistemas de información más comprehensivos, granulares y actualizados. Las instituciones públicas tienen el desafío de actualizarse, transformarse y modernizarse para poder incorporar y aprovechar estos progresos, así como para capitalizar las potenciales interacciones y asociaciones entre el sector privado y público que puedan surgir en este escenario.

REFERENCIAS

- Acemoglu, D. y Autor, D. 2011. Skills, tasks, and technologies: Implications for employment and earnings. In Handbook of labor economics. Vol. 4, pp. 1043-1171. Elsevier.
- Adalet McGowan, M. y D. Andrews. 2015. Skill Mismatch and Public Policy in OECD Countries. OECD Economics Department Working Papers, No. 1210, OECD Publishing, Paris. <https://www.oecd.org/economy/growth/Skill-mismatch-and-public-policy-in-OECD-countries.pdf>
- Altamirano, A. y Amaral, N. 2020. A Skills Taxonomy for LAC: Lessons Learned and a Roadmap for Future Users. Technical Note IDB-TN-02072. Banco Interamericano de Desarrollo. <http://dx.doi.org/10.18235/0002898>
- Altamirano, A., Azuara, O., González, S., Ospina, C., Sánchez, D. y Torres, J. 2019. Clasificación de ocupaciones en América Latina y el Caribe. Technical Note IDB-TN-1832. Banco Interamericano de Desarrollo. <http://dx.doi.org/10.18235/0002062>
- Amaral, N., Eng, N., Ospino, C., Pagés, C., Rucci, G. y Williams, N. 2018. ¿Hasta dónde pueden llevarte tus habilidades? Cómo utilizar los datos masivos para entender los cambios en el mercado laboral. Technical Note IDB-TN-1501. Banco Interamericano de Desarrollo. <http://dx.doi.org/10.18235/0001291>
- Amaral, N., Fieldsend, G., Prada, M., y Rucci, G. 2017. Building Better Skills Systems for Productivity and Growth. Interamerican Development Bank, Labor Markets and Social Security Division. Technical Note No. IDB-TN-1328. Banco Interamericano de Desarrollo. <https://publications.iadb.org/en/building-better-skills-systems-productivity-and-growth>
- Arraño, E. y Jara, K. 2019. Índice de Avisos Laborales de Internet. Estudios Económicos Estadísticos N.º129 Julio 2019, Banco Central de Chile <https://si2.bcentral.cl/public/pdf/estudios-economicos-estadisticos/pdf/see129.pdf>
- Autor, D. H., Levy, F., y Murnane, R. J. 2003. The skill content of recent technological change: An empirical exploration. The Quarterly Journal of Economics, 118(4), 1279-1333.
- Banco Interamericano de Desarrollo. 2020. Skills development sector framework document. Document of the Inter-American Development Bank.

- BGT. 2019. Mapping the genome of jobs. The Burning Glass skills taxonomy. Burning Glass Technologies. https://www.burning-glass.com/wp-content/uploads/2019/09/Burning_Glass_Skills_Taxonomy.pdf
- Busso, M., Bassi, M., Urzúa, S., y Vargas, J. 2012. Desconectados: habilidades, educación y empleo en América Latina. Inter-American Development Bank. <https://publications.iadb.org/es/desconectados-habilidades-educacion-y-empleo-en-america-latina>
- Busso, M., Cristia, J., Hincapié, D., Messina, J., y Ripani, L. 2017. Aprender mejor: políticas públicas para el desarrollo de habilidades. Washington D.C.: Inter-American Development Bank. <http://dx.doi.org/10.18235/0000799>
- Cammeraat, E., y Squicciarini, M. 2021. Burning Glass Technologies' data use in policy-relevant analysis: An occupation-level assessment. OECD Science, Technology, and Industry Working Papers. <https://doi.org/10.1787/cd75c3e7-en>
- Cappelli, P. H. 2015. Skill gaps, skill shortages, and skill mismatches: Evidence and arguments for the United States. ILR review, 68(2), 251-290.
- Cedefop. 2019. Online job vacancies and skills analysis: a Cedefop pan-European approach. Luxembourg: Publications Office. https://www.cedefop.europa.eu/files/4172_en.pdf
- Cedefop. 2021. Understanding technological change and skill needs: big data and artificial intelligence methods. Cedefop practical guide 2. Luxembourg: Publications Office. <https://www.cedefop.europa.eu/en/publications/4198>
- Comyn, P., y Strietska-Ilina, O. (Eds.). 2019. Skills and Jobs Mismatches in Low- and Middle-Income Countries. International Labor Office. https://www.ilo.org/skills/pubs/WCMS_726816/lang--en/index.htm
- Di Capua, L., Queijo, V. y Rucci, G. 2020. Demanda de trabajo en Uruguay: un análisis de vacantes online. Technical Note IDB-TN-1952. Banco Interamericano de Desarrollo. <http://dx.doi.org/10.18235/0002511>
- Gontero, S. y Albornoz, S. 2019. La identificación y anticipación de brechas y habilidades laborales en América Latina: experiencias y lecciones, Serie Macroeconomía del Desarrollo, N° 199. Santiago, Comisión Económica para América Latina y el Caribe (CEPAL). <https://repositorio.cepal.org/handle/11362/44437>
- Gontero, S. y Menéndez, E. 2021. Macrodatos (Big Data) y mercado laboral: identificación de habilidades a través de vacantes de empleo en línea, Documentos de Proyectos (LC/TS.2021/183), Santiago. <https://repositorio.cepal.org/handle/11362/47644>

- Gontero, S. y Zambrano, M. J. 2018. La construcción de sistemas de información sobre el mercado laboral en América Latina. CEPAL. Serie Macroeconomía del Desarrollo, N° 193. Santiago, <https://repositorio.cepal.org/handle/11362/43413>
- González-Velosa, C. y Peña, N. 2019. Demanda de trabajo en América Latina: ¿qué podemos aprender de los portales de vacantes online? Technical Note IDB-TN-1769. Banco Interamericano de Desarrollo. <http://dx.doi.org/10.18235/0001965>
- González-Velosa, C., y Rucci, G. 2016. Métodos para anticipar demandas de habilidades. Technical Note IDB-TN-954. Banco Interamericano de Desarrollo. <https://publications.iadb.org/es/publicacion/17075/metodos-para-anticipar-demandas-de-habilidades>
- Hilbert, M. y Lu, K. 2020. The online job market trace in Latin America and the Caribbean. CEPAL. Disponible en <https://www.cepal.org/en/publications/45892-online-job-market-tracelatin-america-and-caribbean>
- Kureková L, Beblavy M, Thum A. 2014. Using internet data to analyse the labour market: a methodological enquiry, IZA Discussion Paper 8555. IZA, Bonn. <https://docs.iza.org/dp8555.pdf>
- Mazza, J. 2016. Labor Intermediation Services in Developing Economies: Adapting Employment Services for a Global Age. Palgrave McMillan US.
- McGuinness, S., Pouliakas, K., y Redmond, P. 2018. Skills mismatch: Concepts, measurement, and policy approaches. Journal of Economic Surveys, 32(4), 985-1015.
- Mezzanzanica, M., y Mercorio, F. 2019. Big data for labour market intelligence: An introductory guide. European Training Foundation. <https://www.etf.europa.eu/sites/default/files/2019-06/Big%20data%20for%20LMI.pdf>
- Montt, G. 2015. The causes and consequences of field-of-study mismatch: An analysis using PIAAC. OECD Social, Employment and Migration Working Papers, No. 167, OECD Publishing, Paris. <https://doi.org/10.1787/1815199X>
- OECD. 2014. OECD Employment Outlook 2014, OECD Publishing, Paris. OECD. <https://doi.org/10.1787/19991266>
- OECD. 2018. Skills for Jobs. OECD Publishing. www.oecdskillsforjobsdatabase.org
- OIT y OECD. 2018. Approaches to anticipating skills for the future of work. Report prepared by the ILO and OECD for the G20 Employment Working Group. 2nd Meeting of the Employment Working Group Geneva, Switzerland. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms_646143.pdf

- OIT. 2019. Skills for a greener future: A global view based on 32 country studies. International Labour Organization. Geneva: ILO. https://www.ilo.org/wcmsp5/groups/public/---ed_emp/documents/publication/wcms_732214.pdf
- OIT. 2020. The feasibility of using big data in anticipating and matching skills needs – International Labour Organization – Geneva: ILO. https://www.ilo.org/wcmsp5/groups/public/---ed_emp/---emp_ent/documents/publication/wcms_759330.pdf
- OIT/Cinterfor. 2022. Tres ideas derivadas de tres estudios de big data en el mercado de trabajo. OIT/Cinterfor Notas N° 11. <https://www.oitcinterfor.org/tres-ideas-derivadas-tres-estudios-big-data-mercado-trabajo-oitcinterfor-notas-n%C2%BA-11>
- Prada, M. F. y Rucci, G. 2016. Guide to workforce skills assessment instruments. IDB Technical Note IDB-TN-1070. Washington, D.C.: Inter-American Development Bank. <https://publications.iadb.org/en/guide-workforce-skills-assessment-instruments>
- Sattinger, M. 1993. Assignment models of the distribution of earnings. Journal of economic literature, 31(2), 831-880.
- Van Breugel, G. 2017. Identification and anticipation of skill requirements: Instruments used by international institutions and developed countries. Economic Commission for Latin America and the Caribbean (ECLAC). <https://repositorio.cepal.org/handle/11362/42233>
- Vandeplass, A., y Thum-Thysen, A. 2019. Skills Mismatch & Productivity in the EU. Publications Office of the European Union. https://economy-finance.ec.europa.eu/publications/skills-mismatch-and-productivity-eu_en
- Wilson, R. y Zukersteinova, A. 2011. Anticipating changing skill needs: A Master Class. Institute for Employment Research. University of Warwick.

