

Veronika Vilgis
Valeria Jordán
Alejandro Patiño

Medición de la economía de Internet en América Latina

Los casos del Brasil, Chile, Colombia y México



NACIONES UNIDAS

CEPAL



años

Trabajando por
un futuro productivo,
inclusivo y sostenible



BigDATA
Economía digital para
América Latina y el Caribe



DESARROLLO en transición
Instrumento regional
de la Unión Europea para
América Latina y el Caribe

Gracias por su interés en esta publicación de la CEPAL



Si desea recibir información oportuna sobre nuestros productos editoriales y actividades, le invitamos a registrarse. Podrá definir sus áreas de interés y acceder a nuestros productos en otros formatos.

Deseo registrarme



NACIONES UNIDAS



www.cepal.org/es/publications



www.instagram.com/publicacionesdelacepal



www.facebook.com/publicacionesdelacepal



www.issuu.com/publicacionescepal/stacks



www.cepal.org/es/publicaciones/apps

Medición de la economía de Internet en América Latina

Los casos del Brasil, Chile,
Colombia y México

Veronika Vilgis
Valeria Jordán
Alejandro Patiño



Este informe fue elaborado por Veronika Vilgis, Consultora de la División de Desarrollo Productivo y Empresarial de la Comisión Económica para América Latina y el Caribe (CEPAL), y Valeria Jordán y Alejandro Patiño, funcionarios de la misma División. El documento también contó con la colaboración de los Consultores Andrés Garrido y Sebastián Santana. Se agradece el apoyo de Laura Póveda y Valeria Lobos de la CEPAL.

Para la elaboración de este análisis se contó con la colaboración del Instituto Brasileño de Geografía y Estadística (IBGE) y el Centro Regional de Estudios para el Desarrollo de la Sociedad de la Información (CETIC.br) del Brasil, el Instituto Nacional de Estadísticas (INE) de Chile, el Departamento Administrativo Nacional de Estadística (DANE) de Colombia, el Instituto Nacional de Estadística y Geografía (INEGI) de México y NIC Chile de la Universidad de Chile.

Esta publicación se realizó en el marco del proyecto de la Cuenta de las Naciones Unidas para el Desarrollo, denominado “*Big data* para medir y fomentar la economía digital en América Latina y el Caribe”. Asimismo, para esta publicación se contó con el apoyo del programa Mecanismo Regional para el Desarrollo en Transición, de la CEPAL y la Unión Europea, y el proyecto “Observatorio regional de desarrollo digital”.

Ni la Unión Europea ni ninguna persona que actúe en su nombre es responsable del uso que pueda hacerse de la información contenida en esta publicación. Los puntos de vista expresados en este estudio son de los autores y no reflejan necesariamente los puntos de vista de la Unión Europea.

Las opiniones expresadas en este documento, que no ha sido sometido a revisión editorial, son de exclusiva responsabilidad de los autores y pueden no coincidir con las de las Naciones Unidas o las de los países que representan.

Publicación de las Naciones Unidas
LC/TS.2023/51
Distribución: L
Copyright © Naciones Unidas, 2023
Todos los derechos reservados
Impreso en Naciones Unidas, Santiago
S.23-00394

Esta publicación debe citarse como: V. Vilgis, V. Jordán y A. Patiño, “Medición de la economía de Internet en América Latina: los casos del Brasil, Chile, Colombia y México”, *Documentos de Proyectos* (LC/TS.2023/51), Santiago, Comisión Económica para América Latina y el Caribe (CEPAL), 2023.

La autorización para reproducir total o parcialmente esta obra debe solicitarse a la Comisión Económica para América Latina y el Caribe (CEPAL), División de Documentos y Publicaciones, publicaciones.cepal@un.org. Los Estados Miembros de las Naciones Unidas y sus instituciones gubernamentales pueden reproducir esta obra sin autorización previa. Solo se les solicita que mencionen la fuente e informen a la CEPAL de tal reproducción.

Introducción	5
I. La importancia de la transformación digital para el desarrollo sostenible y su medición	7
A. ¿Por qué medir la adopción de tecnologías digitales?.....	7
B. El concepto de economía digital y sus distintas dimensiones.....	7
C. Los retos en la medición de las tecnologías digitales y la importancia de la innovación estadística.....	9
1. <i>Big data</i> : concepto y desafíos metodológicos.....	10
II. Caracterización del Internet de Brasil, Chile, Colombia y México	13
A. Descripción general del panorama de sitios web.....	13
B. El cambiante panorama de Internet: efectos de la pandemia del COVID-1.....	17
III. Analizando la economía de Internet	21
A. Clasificación empresarial según la actividad en línea y definición de la economía de Internet.....	21
B. Composición de la economía de Internet a partir de datos web.....	23
C. La economía de Internet en Brasil, Chile, Colombia y México: resultados de la combinación de datos web y registros administrativos.....	26
1. Efectos de la pandemia: seguimiento de los cambios entre marzo y septiembre de 2020.....	31
2. Consideraciones metodológicas en la combinación de datos web y registros empresariales	33
3. Características de los sitios web que fueron vinculados con el registro empresarial.....	35
IV. Reflexiones finales	37
Bibliografía	39
Cuadro	
Cuadro 1 Cantidad de empresas por categoría según actividad en línea para cada país, 2020.....	26
Gráficos	
Gráfico 1 Sitios web disponibles y no disponibles en Brasil, Chile, Colombia y México al mes de septiembre de 2019, 2020 y 2021.....	13
Gráfico 2 Distribución de los sitios web disponibles según tipo a septiembre de 2019, 2020 y 2021.....	14
Gráfico 3 Distribución de los sitios web disponibles de cada país según el nombre de dominio a septiembre de 2019, 2020 y 2021.....	15
Gráfico 4 Distribución de los sitios web disponibles según ubicación de alojamiento a septiembre de 2019, 2020 y 2021.....	16
Gráfico 5 Sitios web disponibles con protocolo IPV6, a septiembre de cada año.....	17
Gráfico 6 Proporción de sitios web etiquetados como “sensibles a la privacidad”, septiembre de 2021	17
Gráfico 7 Nuevos sitios web empresariales encontrados entre febrero y agosto de 2019, 2020 y 2021.....	18
Gráfico 8 Distribución de nuevos sitios web empresariales por rubro de actividad económica en Brasil, Colombia, México y Chile, entre marzo 2020 y octubre 2021.....	19
Gráfico 9 Porcentaje de sitios web empresariales con al menos un método de pago disponible en 2019, 2020 y 2021	19
Gráfico 10 Distribución de sitios web empresariales según tipo de actividad en línea por país, septiembre 2021.....	23
Gráfico 11 Huella económica de los sitios web empresariales según el tipo de actividad en línea, septiembre de 2021.....	24

Gráfico 12	Variación de la cantidad de sitios web empresariales por tipo de actividad en línea entre marzo de 2020 y septiembre de 2021.....	25
Gráfico 13	Sitios web empresariales nuevos y que cambiaron de categoría entre marzo y septiembre de 2020, y entre septiembre de 2020 y septiembre de 2021 en los cuatro países.....	25
Gráfico 14	Variación en la cantidad de sitios web empresariales según categoría entre los periodos marzo-septiembre de 2020 y septiembre de 2020-septiembre de 2021 en los cuatro países.....	26
Gráfico 15	Porcentaje de empresas vinculadas exitosamente con un sitio web por categoría, 2020.....	27
Gráfico 16	Participación de la industria (código SIC) por categoría de internet.....	28
Gráfico 17	Participación de las empresas por categoría presencia en Internet según su tamaño (número de empleados).....	29
Gráfico 18	Participación de las empresas según su volumen de ventas anuales por categoría de presencia en línea en Chile, marzo de 2020.....	31
Gráfico 19	Participación de empresas identificadas como exportadoras, importadoras o ambas, Chile, marzo de 2020.....	31
Gráfico 20	Empresas que cambiaron su tipo de presencia en línea	32
Gráfico 21	Disponibilidad información en la base de datos web para seis variables utilizadas para la combinación (información a noviembre de 2021).....	34
Gráfico 22	Disponibilidad de información dentro del registro de empresas de cada país para cinco variables utilizadas para la combinación.....	34
Gráfico 23	Huella económica promedio de los sitios web que se fusionaron y los que no se fusionaron en los registros comerciales nacionales.....	35
Gráfico 24	Porcentaje de sitios web vinculados y no vinculados según perfiles de redes sociales listados en el sitio web.....	36

Mapa

Mapa 1	Diferencias regionales de empresas sin sitio web para Brasil y Colombia.....	30
--------	------------------------------------------------------------------------------	----

Diagramas

Diagrama 1	Dimensiones y elementos del desarrollo digital	9
Diagrama 2	Clasificación empresarial según su actividad en línea	22

Las tecnologías digitales son una herramienta fundamental para la diversificación y sofisticación de los sectores productivos, y la inclusión social. Por este motivo, la adopción de estas tecnologías es un factor crítico y esencial para el funcionamiento de la economía y la sociedad. América Latina y el Caribe afronta el reto de impulsar la transformación digital para dar respuesta a desafíos estructurales como el bajo crecimiento económico, los bajos niveles de inversión y los crecientes niveles de pobreza, además de reducir las brechas de productividad con países más avanzados. Para esto se deben diseñar políticas basadas en evidencia que puedan impulsar la innovación y el cambio tecnológico.

Conforme avanza la digitalización de actividades económicas y sociales, la disponibilidad de datos crece exponencialmente habilitando nuevas fuentes de información, que en combinación con herramientas novedosas impulsan la ciencia y la innovación de datos. Esta nueva información puede convertirse en un insumo clave para comprender la dinámica de la economía digital y mejorar el diseño de políticas que impulsen estas transformaciones.

La digitalización se aceleró debido a la pandemia del COVID-19, impulsando el teletrabajo, la educación a distancia, y las actividades en Internet. En particular, la proliferación de tiendas en línea y soluciones de comercio electrónico se consolidaron en este periodo. En este contexto, las estadísticas tradicionales disponibles no lograron proporcionar información oportuna sobre estos cambios (CEPAL y Adenauer, 2021). Por lo tanto, explorar nuevos métodos de medición puede ser una alternativa relevante para comprender mejor fenómenos emergentes.

El presente estudio se desarrolló como un ejercicio exploratorio que combina fuentes de datos extraídos de la web con fuentes estadísticas oficiales, específicamente los registros empresariales de las Oficinas Nacionales de Estadística (ONE), con el objetivo de dimensionar las actividades empresariales en línea, sin restringirse a clasificaciones industriales tradicionales. El estudio busca fomentar la innovación de datos y la exploración del uso de técnicas y herramientas de macrodatos (*big data*) para aportar aspectos metodológicos y conceptuales que puedan mejorar la comprensión de estos fenómenos, y servir de base para futuras investigaciones.

El trabajo realizado permitió tener una clasificación novedosa de empresas en función al uso que realizan de Internet, y que conforman la economía de Internet. Esta caracterización se realizó en cuatro países de América Latina: Brasil, Colombia, Chile y México, entre 2019 y 2021. La primera parte del documento analiza la importancia y los retos de la medición de la economía digital, presentando también definiciones conceptuales relevantes para el estudio. Posteriormente, se presenta un análisis de sitios web empresariales para los cuatro países del estudio. En la última parte del reporte, se revisan los resultados de la combinación de datos web con los datos de empresas de los registros comerciales, incluyendo las consideraciones metodológicas correspondientes. El documento finaliza con una reflexión donde se repasan los principales hallazgos del estudio, pero también los retos encontrados.

A. ¿Por qué medir la adopción de tecnologías digitales?

La revolución digital está dando lugar a una nueva era industrial basada en las tecnologías digitales, donde la economía tradicional se superpone y fusiona con sus esquemas organizativos, productivos y de gobernanza con la economía digital y sus innovaciones, modelos de negocio, de producción y de organización empresarial (CEPAL, 2018; CEPAL, 2022)

A medida que la adopción de las tecnologías digitales se impregna en el tejido económico y social, se evidencia el rol clave de la digitalización para el desarrollo de los países por su potencial de abrir un camino a sociedades con mayor bienestar y economías más productivas y competitivas. Esto es particularmente relevante para países condicionados por brechas estructurales, como es el caso de los de América Latina y el Caribe. Sin embargo, la transformación digital no es un proceso lineal y requiere de un conjunto de políticas con un enfoque estratégico, que involucre diversos ámbitos de la política, como el desarrollo de infraestructura, la promoción de habilidades, el entorno regulatorio y normativo y la promoción de la innovación, por citar algunos. Frente a este escenario es necesario contar con información estadística confiable y comparable que permita diseñar, monitorear y evaluar políticas y regulaciones basadas en evidencia.

La comunidad estadística internacional, las ONE y diversos organismos internacionales han realizado avances importantes en generar referencias y marcos estadísticos estandarizados para la producción de indicadores sobre el grado de adopción de tecnologías de la información y las comunicaciones (TIC), principalmente en materia de conectividad a Internet por parte de las personas, los hogares y las empresas. No obstante, y a pesar de los avances logrados, los métodos de medición no han logrado mantenerse al ritmo de la evolución de las tecnologías digitales y su permeabilidad en el tejido social y el productivo. Los desafíos son diversos e implican la actualización de métricas e indicadores sobre la adopción de tecnologías digitales emergentes, la obtención de información de forma más granular y oportuna, y la observación de una nueva generación de usos, actores y modelos de negocio.

B. El concepto de economía digital y sus distintas dimensiones

No existe una definición única o universalmente acordada de la economía digital. Esto se debe en parte al carácter multidimensional que tiene este concepto y su versatilidad. No obstante, un número importante de organismos vinculan esta definición con aquella economía que sustenta sus procesos y capacidad de creación de valor en tecnologías digitales, o como una forma de describir como estas tecnologías están cambiando los patrones de consumo y de producción.

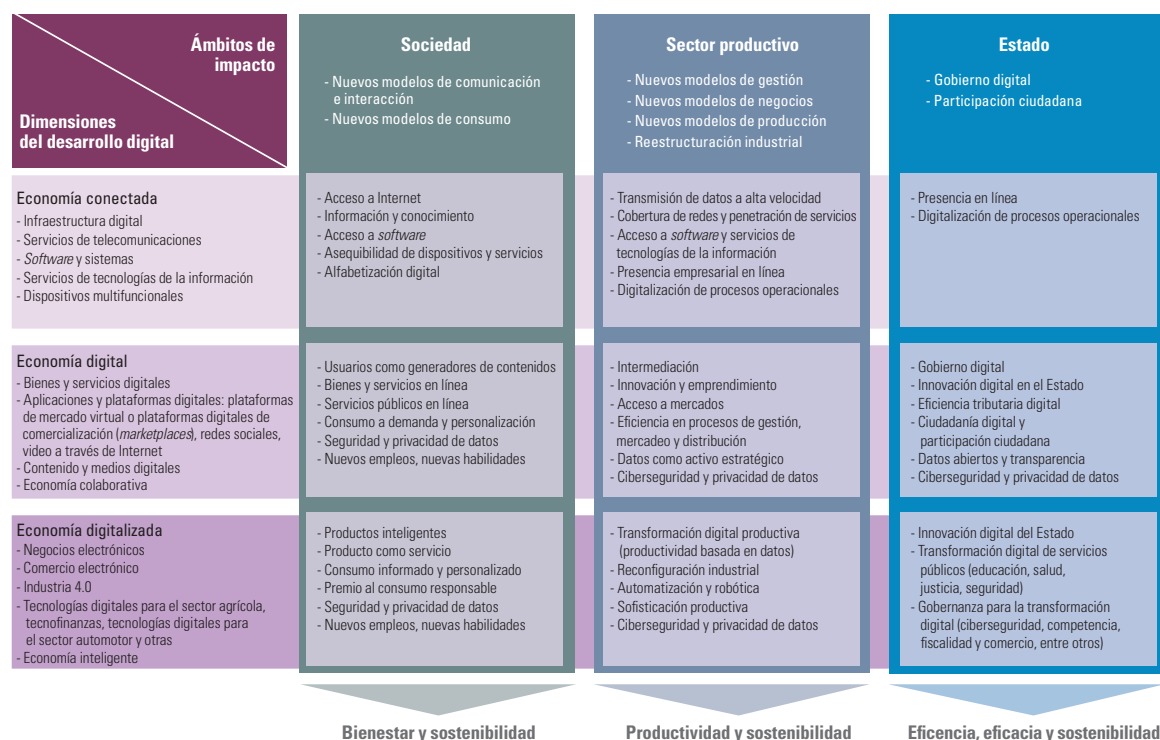
Según la Comisión Europea (2014) la economía digital es una “economía basada en tecnologías digitales (a veces denominada economía de Internet)”. Por otra parte, en su definición más reciente la Organización para la Cooperación y el Desarrollo Económico (OECD, 2020) define este término como “toda actividad económica que depende del uso de insumos digitales o cuyo desempeño mejora significativamente mediante su uso, incluyendo las tecnologías digitales, la infraestructura digital, los servicios digitales y los datos. Considerando a todos los productores y consumidores, incluido el gobierno, quienes utilizan estos insumos digitales en sus actividades económicas”. La Conferencia de las Naciones Unidas sobre Comercio y Desarrollo (UNCTAD, 2022) retoma el concepto de la OCDE y enfatiza la importancia de tener una aproximación flexible que permita analizar diferentes componentes de la economía digital como los productos (bienes y servicios de TIC, contenidos digitales, servicios prestados digitalmente), las transacciones (compras o entregas digitales) y la producción (grado de dependencia de insumos digitales).

Por otro lado, la CEPAL (2021) plantea un marco de análisis de tres dimensiones para estudiar el potencial de las tecnologías digitales para aumentar el bienestar de las personas, la productividad de las empresas, la eficiencia y la eficacia del Estado y la sostenibilidad medioambiental. Estas dimensiones son la economía conectada, la economía digital y la economía digitalizada (véase diagrama 1):

- Economía conectada: corresponde a la infraestructura digital (redes de banda ancha, puntos de intercambio de tráfico, centros de datos, etc.), la masificación de diversos dispositivos (computadores, laptops, tabletas y teléfonos inteligentes) para el uso de Internet, y el uso de sensores que habilitan la conexión no sólo de las personas, sino que también de las máquinas y objetos.
- Economía digital: se compone de la parte de la producción económica derivada principalmente de modelos de negocios basados en tecnologías digitales para la oferta de bienes y servicios. En esta dimensión se incluyen los modelos de plataformas e intermediarios digitales que cambian la propuesta de valor en diversos sectores económicos¹, mediante el uso intensivo de datos, la optimización de los procesos operativos, la segmentación de mercados y la personalización de productos y servicios.
- Economía digitalizada: se trata de la transformación de los modelos de negocios y producción de las empresas de sectores tradicionales en sistemas inteligentes conectados, gracias a la adopción de tecnologías avanzadas, como las redes móviles de quinta generación (5G), Internet de las cosas, la computación en la nube, la inteligencia artificial, la realidad virtual y aumentada, el análisis de macrodatos y la robótica cognitiva. El objetivo es incrementar la competitividad, productividad y sostenibilidad de las operaciones.

¹ En la actualidad las plataformas digitales operan en múltiples sectores económicos, facilitando las transacciones comerciales (Amazon, Alibaba o Mercado Libre), los servicios financieros (Ant Group, Avant, Mercado Pago o Nubank), las comunicaciones y las redes sociales (Facebook, Instagram o WhatsApp), el turismo y el hospedaje (Despegar, Booking o Airbnb), el desarrollo de aplicaciones (IOS de Apple y Android de Google), el encuentro de oferta y demanda de empleo (Laborum, LinkedIn, Workana o Freelancer), entre otras.

Diagrama 1
Dimensiones y elementos del desarrollo digital



Fuente: Comisión Económica para América Latina y el Caribe (CEPAL), *La Inversión Extranjera Directa en América Latina y el Caribe, 2021* (LC/PUB.2021/8-P), Santiago, 2021.

C. Los retos en la medición de las tecnologías digitales y la importancia de la innovación estadística

Tradicionalmente, la medición estadística de la economía digital se ha abordado desde el lado de la producción (oferta) o del consumo (demanda) de las TIC, así como desde el despliegue de la infraestructura y servicios de telecomunicaciones y el comercio de productos TIC. Por ejemplo, desde el lado de la oferta se pueden generar indicadores sobre las industrias y empresas que proveen infraestructura, bienes y servicios TIC, mientras que por el lado de la demanda se puede analizar los niveles de adopción y uso de TIC por parte de empresas y las personas. Parte del diseño de este tipo de estadísticas guarda relación con la definición del sector TIC, que abarca principalmente las actividades de la manufactura, el comercio y los servicios TIC. A partir de este diseño se pueden realizar estimaciones sobre la contribución de las TIC al crecimiento y el empleo (UNCTAD, 2022).

Si bien existe una definición del sector TIC², esta dimensión no logra capturar todas las facetas y perspectivas de la economía digital³ en la actualidad. Por otro lado, no existe una definición estadística acordada sobre este concepto, por lo que las oficinas nacionales de estadística a nivel internacional utilizan diferentes aproximaciones. Frente a

² (OECD, 2011) usando la Clasificación Industrial Internacional Uniforme (CIIU) Rev. 4.

³ En esta clasificación se excluyen empresas que hacen un uso intensivo de tecnologías digitales como las plataformas, las empresas de comercio electrónico o los servicios digitales. Existen también otras definiciones como las industrias de la información que además del sector TIC incorporan los sectores de contenido audiovisual.

este vacío, la Hoja de Ruta del G20 sobre un marco común de medición de la economía digital, propone una clasificación por niveles. El primer nivel supone una “medición básica” de la economía digital donde solamente se incluyen las actividades económicas derivadas de proveedores de bienes y servicios TIC. El segundo nivel incorpora una “medición reducida” donde a los productos de bienes y servicios TIC se suma la actividad económica derivada de empresas que dependen de insumos digitales. El tercer nivel se considera una “medición amplia”, pues a las primeras dos dimensiones, se suma la actividad económica de empresas que mejoran significativamente su desempeño mediante insumos digitales. Finalmente, esta clasificación también hace referencia a la sociedad digital para incluir en el análisis el resto de los efectos de las tecnologías digitales en la sociedad y la economía (OCDE, 2020).

La clasificación provista en la Hoja de Ruta del G20 permite tener un marco de análisis más preciso de la economía digital, sin embargo, el mismo documento advierte que todavía existen ambigüedades al momento de definir el nivel de dependencia de insumos digitales de las empresas (básica, reducida y/o amplia), para lo cual sugiere continuar trabajando en lineamientos o criterios adicionales para ayudar en esta delimitación por niveles.

En este contexto, el presente estudio pretende realizar un aporte al abordar la medición de la economía digital desde una perspectiva innovadora que permita clasificar a las empresas según el nivel de dependencia de Internet en su modelo de negocios, utilizando técnicas de grandes datos para extraer información de la web y su combinación con otras estadísticas tradicionales, y así poder dimensionar, lo que se denomina la “economía de Internet”. Esta aproximación permite analizar segmentos clave de empresas que basan todo su modelo de negocio en Internet, cómo plataformas de comercio electrónico, servicios profesionales y aplicaciones, pero también dimensionar el nivel de uso de Internet del resto de empresas de la economía.

1. *Big data*: concepto y desafíos metodológicos

El rápido progreso tecnológico impulsado por la revolución digital ha provocado un aumento exponencial de la capacidad de recopilar, almacenar, procesar y analizar datos. Las actividades realizadas en línea dejan un rastro digital que produce grandes cantidades de datos que se han convertido en un bien intangible y que permiten desarrollar nuevas tecnologías disruptivas como la inteligencia artificial.

El término “*big data*” se ha acuñado para referirse al gran volumen de información que produce la huella digital. Sin embargo, el tamaño por sí solo no define al *big data*, sino que influyen otros aspectos, tales como la variedad y la velocidad en la que se producen y procesan datos (Layney, 2001). La variedad implica la expansión de los datos estructurados tradicionales (tabulares) a datos no estructurados, como texto, imágenes, información de video y audio. La velocidad se refiere a la rapidez a la que se producen y analizan los datos, lo que a menudo permite el procesamiento en tiempo real de diversas transacciones. Desde la introducción de estas tres dimensiones o tres “V”, se han añadido también otras características al término *big data*, incluyendo la veracidad (calidad de datos), la variabilidad (cambios en la velocidad a la que se generan los datos) y el valor (los datos en su forma original a menudo tienen un valor bajo en relación con su volumen).

Sin embargo, las descripciones anteriores solo se centran en lo que se entiende por *big data* y no toman en cuenta lo que hace *big data*. Por esta razón, es mejor hacer referencia a *big data* como nuevas técnicas de análisis de datos, y de esta forma tomar en cuenta los métodos analíticos que surgen a partir de esta tendencia. El análisis

de macrodatos requiere herramientas y programas especializados para almacenar y analizar grandes volúmenes de datos no estructurados y altamente variables (Ward y Barker, 2013).

El análisis de *big data* ofrece una oportunidad única para medir aspectos de la economía digital que los enfoques tradicionales de medición no pueden capturar o no pueden proporcionar dentro de un marco de tiempo oportuno para analizar los rápidos cambios que ocurren en este nuevo paradigma. Al tratarse de un fenómeno nuevo y en constante evolución, el seguimiento de la economía digital demanda no solo nuevas métricas que den cuenta de la dinámica de los diferentes elementos que la componen, sino también de innovar en las técnicas de medición. Así, a medida que las actividades socioeconómicas dejan cada vez más rastros digitales, la analítica y la ciencia de datos ofrece nuevas oportunidades.

El desafío que ya no radica en la disponibilidad de datos, sino en la capacidad de usar estas nuevas herramientas, lo que depende de algunos elementos habilitadores como la infraestructura tecnológica, las habilidades técnicas y la capacidad de extraer significado de los datos. Por otra parte, también existen ciertos aspectos que merecen atención al momento de usar estas herramientas. Los datos que produce la huella digital son orgánicos (transaccionales y no estructurados) y no son producto de un diseño estadístico determinado (ej. encuestas), tampoco surgen de muestras aleatorias, lo que dificulta establecer patrones generales para aplicarlos a toda la población. Por otra parte, en algunas ocasiones estos datos provienen de servicios específicos como las plataformas⁴ o la Internet, y por ende mantienen un sesgo al no ser representativos, dado que no toda la población tiene acceso a estos servicios. Adicionalmente, la experiencia ha demostrado que el trabajo de limpieza y armonización que requiere usar *big data* es significativo y requiere de ciertos parámetros y lineamientos, para los cuáles en muchos casos no cuentan con metodologías estandarizadas. De esta forma, todo podría indicar que trabajar con ese tipo de herramientas no reemplazará por completo las estadísticas tradicionales, sino en todo caso las complementarán (CEPAL, 2020).

Otra advertencia crucial con respecto al análisis de *big data* son los resguardos que se debe tener respecto a la privacidad y los datos personales. Estas herramientas, sobre todo la vinculación entre ciencia de datos y el uso de redes sociales, permiten identificar a las personas de forma más fácil e incluso desarrollar sistemas y técnicas para predecir el comportamiento individual (Davahli y otros, 2020; Markovikj y otros, 2013). Como tal, la recopilación y el uso ético de los datos es un imperativo en el uso de *big data*. Con este fin existen manuales que definen principios para guiar la identificación y resolución de problemas éticos que surjan de la investigación que involucren este tipo de fuentes de información y tecnologías, como el reporte de Menlo (Bailey y otros, 2012).

⁴ Ej. Twitter o Uber.

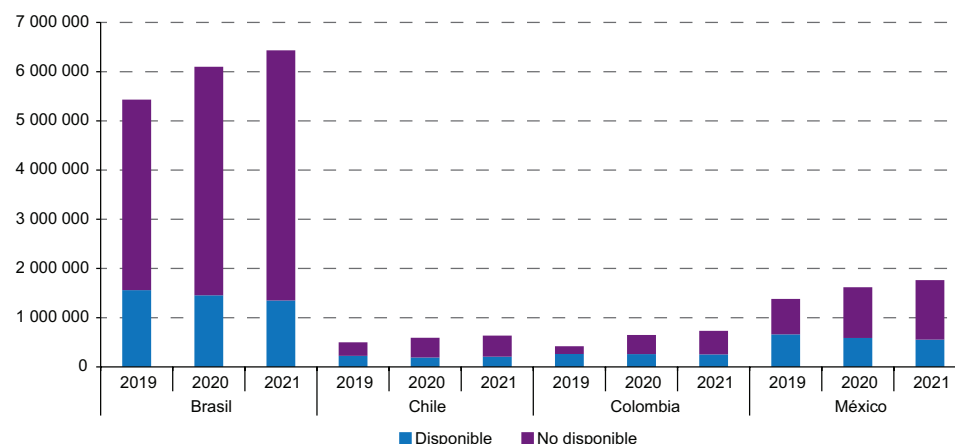
A. Descripción general del panorama de sitios web

A fin de dimensionar la presencia y actividades empresariales en línea, el presente ejercicio parte analizando la información de los sitios web de Brasil, Chile, Colombia y México. Los datos que permiten esbozar el panorama de Internet de cada país son generados mediante el uso de técnicas de raspado web (*web scraping*) con una herramienta que permite indexar más de 450 millones de nombres de dominios en 50 países⁵. A partir de la información recolectada se creó una base de datos con variables que incluyen información del contenido del sitio web, sus características técnicas y datos comerciales. El país de origen del sitio web se determinó examinando un conjunto de variables, como el país de alojamiento del sitio, el idioma, el nombre de dominio y los datos de contacto indicados en el propio sitio web (por ejemplo, códigos postales, números de teléfono, direcciones, etc.). Cabe indicar, que la base de datos incluye tanto los nombres de dominio a nivel país (.cl, .br, .co, .mx) como los nombres de dominio genérico (ej. .com, .org, .net).

En septiembre de 2021, se identificaron 6.434.750 sitios web pertenecientes a Brasil, 636.405, a Chile, 732.918 a Colombia y 1.763.504 a México. Sin embargo, luego de realizar una limpieza de los registros, se eliminaron los sitios “sin respuesta” debido a denegación de acceso o aquellos que mostraban un servidor no encontrado (“*host not found*”), quedando disponibles alrededor del 25% de los sitios. Esto representa un número final de sitios web disponibles de 1.350.163 en Brasil, 208.077 en Chile, 254.441 en Colombia y 554.494 en México, estos son sitios web accesibles al público y que proveen algún tipo de contenido (véase gráfico 1). Cabe destacar, que entre 2019 y 2021, los cuatro países registraron una disminución de los sitios disponibles, -16% en México, -13% en Brasil, -7% en Chile y -3% en Colombia.

Gráfico 1
Sitios web disponibles y no disponibles en Brasil, Chile, Colombia y México al mes de septiembre de 2019, 2020 y 2021

(En cantidad de sitios web)



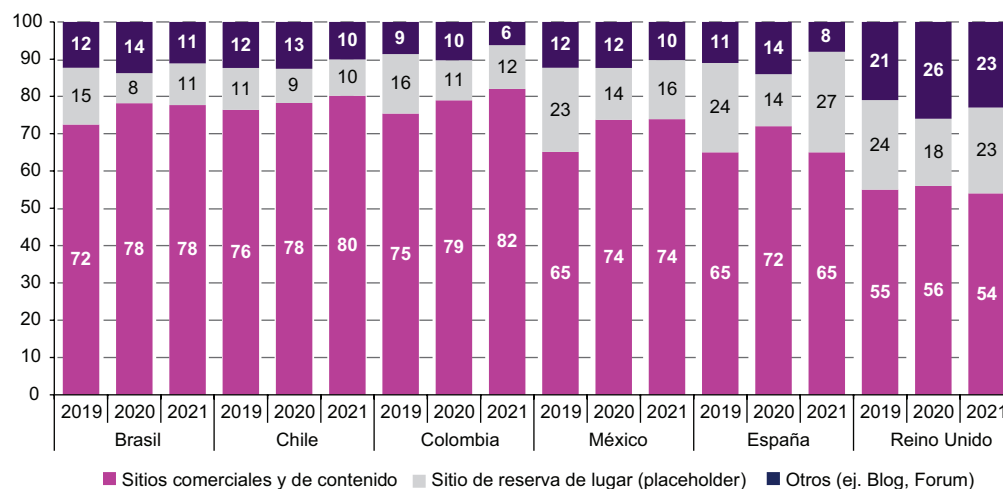
Fuente: CEPAL con datos de Dataprovider.

⁵ Esta herramienta fue provista por la empresa Dataprovider, más información en: Dataprovider.com.

El gráfico 2 presenta la distribución de los sitios web disponibles según su tipo para los cuatro países del estudio, además de España y Reino Unido. Se distinguen tres categorías: (i) sitios web comerciales y de contenido, (ii) sitios web de reserva de lugar (*placeholder*) y (iii) otros (blogs, foros, etc.). La primera categoría incluye a los sitios web empresariales, sitios de tiendas en línea o de comercio-e, y sitios de contenido. Los sitios web empresariales se refieren a páginas con información corporativa sobre los productos o servicios, pudiendo incluir funcionalidades de compra en línea. Los sitios de tiendas en línea se refieren a páginas de empresas que son exclusivamente plataformas de comercialización por medios electrónicos. Los sitios web de contenido por su parte, consideran cualquier elemento creativo, ya sea texto, imágenes, archivos de audio y video, mensajes de correo electrónico archivados, datos, servicios electrónicos (de reserva de hora, calculadoras), aplicaciones, etc; y pueden tener intención comercial, siendo utilizados por proveedores, profesionales independientes, por ejemplo.

Gráfico 2

Distribución de los sitios web disponibles según tipo a septiembre de 2019, 2020 y 2021
(En porcentajes)



Fuente: CEPAL con datos de Dataprovider.

A septiembre de 2021, en los cuatro países, la mayor proporción de los sitios web disponibles era de tipo comercial y de contenido, representando el 82% en Colombia, el 80% en Chile, el 78% en Brasil y el 74% en México. Estos valores fueron más altos que los registrados en España y el Reino Unido, en donde este tipo de sitios representó 65% y 54%, respectivamente. Por otro lado, los sitios de tipo marcador de posición (*placeholder*), que sirven para reservar un sitio web que en algún momento será completado con algún tipo de contenido, representaron entre el 10% (Chile) y 16% (México). Esta categoría contenía alrededor de un cuarto del total de los sitios disponibles de los países europeos de la muestra.

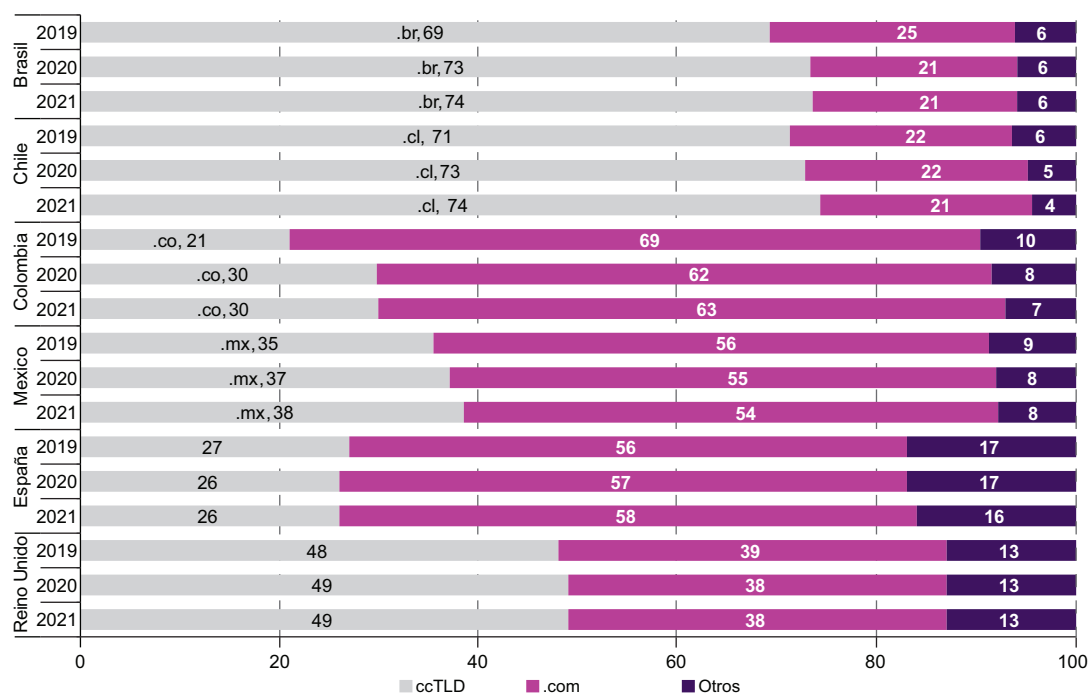
Un análisis temporal entre septiembre de 2019 y septiembre de 2021 muestra que el peso relativo de los sitios comerciales y de contenido aumentó en los países de la región, en 14% en el caso de México, 9% en Colombia, 7% en Brasil y 5% en Chile, en tanto que en España y Reino Unido tal participación prácticamente se mantuvo constante. Esto sugeriría una mayor presencia empresarial en línea durante los meses de pandemia en los países de la región, dada la necesidad de mantener las actividades de comercialización con cuarentenas más restrictivas y duraderas, particularmente en los casos de Chile y Colombia. Como se verá más adelante, tal situación se explica aún

más si se considera que los sitios de tiendas en línea duplicaron su participación en los países de la región, en ese mismo periodo. Por el contrario, en los países europeos de la muestra, dicha participación no varió.

El gráfico 3, muestra la distribución de los sitios web de cada país según el nombre de dominio utilizado, es decir si se usa un dominio de nivel superior de código país (ccTLD, por sus siglas en inglés) o un dominio genérico, por ejemplo, el ".com" (gTLD, por sus siglas en inglés). México y Colombia se destacan por un uso intensivo del dominio genérico (cifras cercanas al 60%), mientras que, por el contrario, en Brasil y Chile el 74% de los sitios web pertenecían al dominio país en 2021. En comparación con países de Europa, se tiene que España también cuenta con una proporción relativamente baja (26%) de sitios que utilizan código país y la proporción de sitios web ".com" alcanzó el 58% en 2021. En el Reino Unido, sucede lo opuesto, dado que alrededor de la mitad de los dominios corresponden a la terminación ".uk". Cabe indicar, que ambos países europeos tienen una diversidad mayor en los gTLD con una participación significativa de otras terminaciones además del ".com" (por ejemplo, .org, .net). Las razones que explican estas diferencias pueden ser diversas, y podrían estar relacionadas con las estrategias comerciales de los registros de dominios y las propias estrategias comerciales de posicionamiento web de las empresas que usan estos dominios.

Gráfico 3
Distribución de los sitios web disponibles de cada país según el nombre de dominio a septiembre de 2019, 2020 y 2021

(En porcentajes)

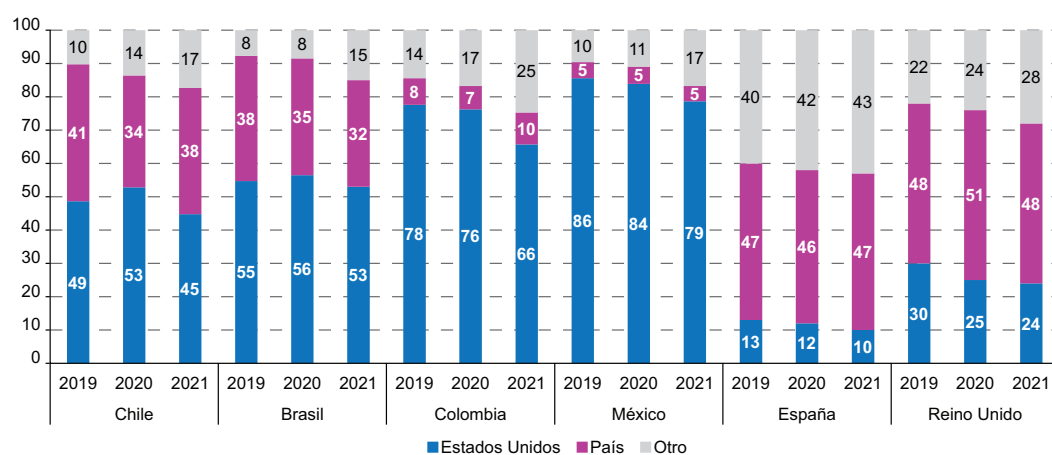


Fuente: CEPAL con datos de Dataprovider.

En cuanto al lugar de alojamiento de los sitios web, se observa una propensión de los cuatro países analizados por el uso de servidores situados en Estados Unidos de Norteamérica, principalmente (véase gráfico 4). Esta proporción es mayor en México y Colombia con más del 65% de sus sitios web alojados en dicho país, según los datos

a septiembre de 2021. No obstante, se observa una reducción significativa desde 2019 cuando representaban cerca del 80% en Colombia y 90% en México. Cabe advertir que en el caso de España y el Reino Unido, la mitad de los sitios web estaban alojados localmente, una proporción significativamente mayor. La ubicación del servidor en el alojamiento web es clave para lograr tiempos de carga más rápidos de los sitios. Esto, se traduce en mejor experiencia de navegación para los clientes, lo que equivale a menos abandonos, mayores probabilidades de convertir a visitantes en clientes y mejores clasificaciones en los motores de búsqueda web, ya que los robots de rastreo asociarán el lugar de alojamiento del sitio con la ubicación geográfica de la empresa.

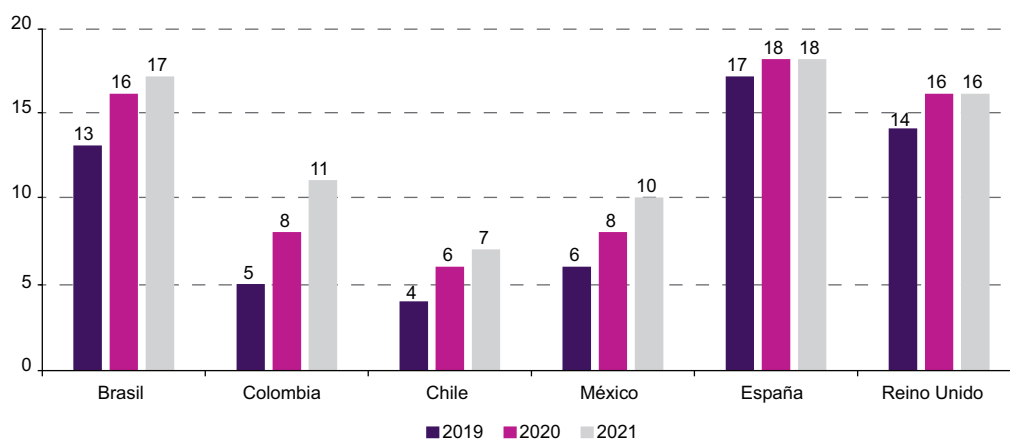
Gráfico 4
Distribución de los sitios web disponibles según ubicación de alojamiento a septiembre de 2019, 2020 y 2021
(En porcentajes)



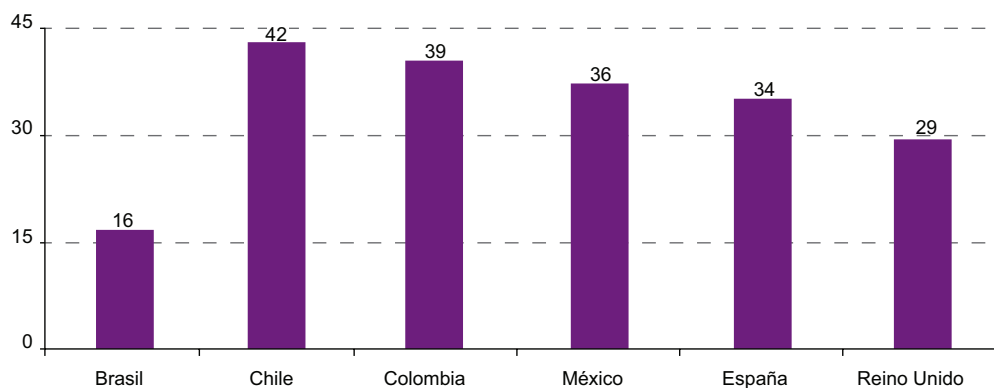
Fuente: CEPAL con datos de Dataprovider.

Ipv6 es la última versión del Protocolo de Internet (IP, por sus siglas en inglés) en el cual se sustenta la operación de Internet. La principal motivación para el diseño y despliegue de Ipv6 sobre su antecesor, el protocolo Ipv4, fue la expansión del espacio de direcciones disponible en Internet y el advenimiento de tendencias como el Internet de las Cosas, además de la necesidad de conectar más dispositivos a la red. La información técnica existente en los sitios web permite evidenciar la creciente adopción del protocolo Ipv6 en los países analizados. El gráfico 5 muestra que, en septiembre de 2021, Brasil lideraba su adopción con el 17% de los sitios web con una dirección Ipv6, seguido por Colombia (11%), México (10%) y Chile (7%). La adopción de Brasil es similar a España (18%) y Reino Unido (16%).

La protección y la privacidad de los datos son aspectos cada vez más importantes, pues condicionan el grado de confianza de soluciones digitales. Entre los cuatro países, Chile tiene la mayor proporción de sitios web sensibles a la privacidad, en el entendido que pueden almacenar información personal en la medida que ofrecen la posibilidad de crear cuentas, iniciar sesión en un sistema, completar un formulario de contacto o procesar un pago (véase gráfico 6).

Gráfico 5**Sitios web disponibles con protocolo IPV6, a septiembre de cada año***(En porcentaje de sitios web disponibles totales)*

Fuente: CEPAL con datos de Dataprovider.

Gráfico 6**Proporción de sitios web etiquetados como "sensibles a la privacidad", septiembre de 2021***(En porcentajes)*

Fuente: CEPAL con datos de Dataprovider.

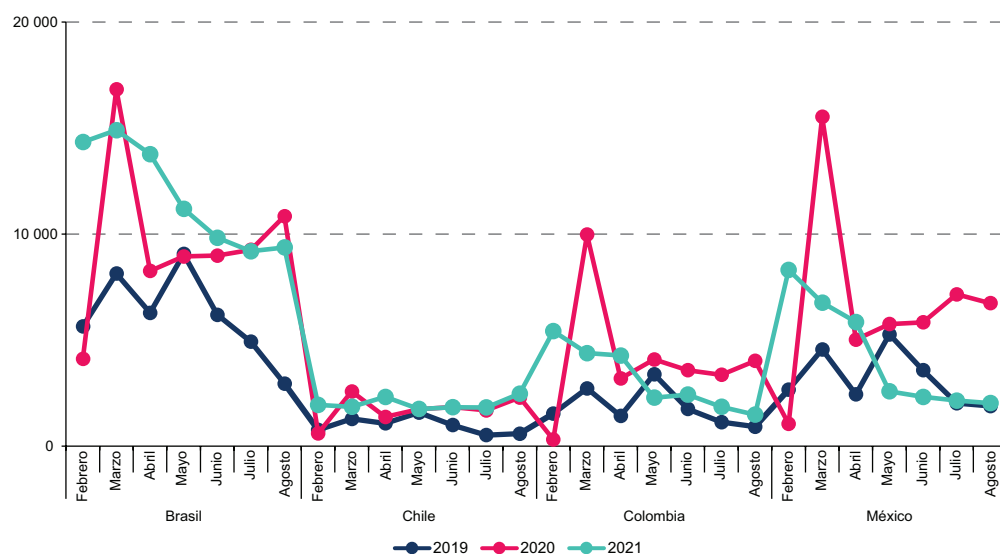
B. El cambiante panorama de Internet: efectos de la pandemia del COVID-19

Quizás ningún otro suceso obligó a tantas personas y empresas a adoptar rápidamente las nuevas herramientas digitales y a trasladar sus actividades en línea como la pandemia del COVID-19. Con el establecimiento de cuarentenas y otras medidas de distanciamiento social impuestas por los gobiernos, muchas tiendas, restaurantes y servicios, se vieron obligados a trasladar sus actividades a Internet. Muchas empresas que hasta el momento sólo tenían un sitio web estático o un perfil de redes sociales y no habían visto la necesidad de tener una presencia activa en línea, se enfrentaron con la urgencia de habilitar canales virtuales de comercialización y relacionamiento con los clientes. Es en este contexto, que el presente estudio es particularmente útil para examinar los cambios producto del COVID-19.

De hecho, al analizar la cantidad de nuevos sitios web empresariales durante los primeros meses de confinamiento en comparación con los meses previos a la pandemia, se observa un aumento significativo en los cuatro países analizados. El mayor crecimiento se registró en marzo de 2020, lo que coincide con las primeras medidas de distanciamiento social. Luego del *boom*, la cantidad de nuevos sitios web empresariales disminuye para posteriormente experimentar un crecimiento sostenido en los siguientes meses de ese año. Así, entre febrero y agosto de 2020, el crecimiento promedio mensual de los sitios web fue de 48% en Brasil, 57% en Chile, 511% en Colombia y 224% en México. Estas altas tasas de crecimiento llaman aún más la atención al compararlas con los mismos meses del año 2019, en los que Brasil experimentaba una disminución del 4% de los sitios web disponibles, y Chile, Colombia y México registraban incrementos de entre 4% y 11%. En 2021, la tendencia al crecimiento de los sitios web se moderó e incluso se revirtió con excepción de Chile, en donde continuaron aumentando en un promedio de 6% mensual y donde se mantenían severas restricciones de movilidad. Entre marzo y agosto de 2021, Brasil y Chile acumularon más sitios web nuevos que en el mismo período en 2020, en tanto que Colombia y México registraron una cantidad menor en 2021 respecto al año anterior, pero mayor a 2019.

Gráfico 7
Nuevos sitios web empresariales encontrados entre febrero y agosto de 2019, 2020 y 2021

(Cantidad de sitios web)



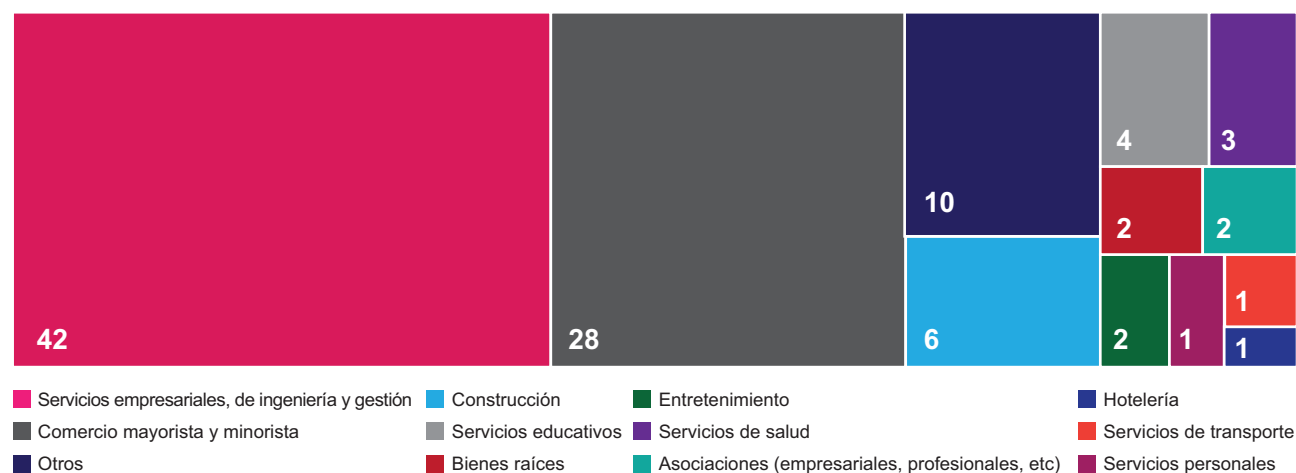
Fuente: CEPAL con datos de Dataprovider.

Cabe destacar, que además del surgimiento de nuevos sitios empresariales, también se observó un cambio en el tipo de presencia en línea entre enero y junio de 2020. En los primeros seis meses de 2020, en el Brasil, Chile, Colombia y México se observó que el 20% de los cambios en los sitios web existentes correspondió a transición a sitios de tipo transaccional (CEPAL, CAF, y Digital Policy Law, 2020).

Cuando se examinan los sitios web comerciales que aparecieron entre marzo de 2020 y septiembre de 2021 según su clasificación por segmento industrial, se observa que los sectores que experimentaron un mayor aumento fueron los servicios empresariales, de ingeniería y de gestión, que representaron el 42% del total de los nuevos sitios empresariales, seguidos de los de comercio mayorista y minorista con el 28%. Destaca también la mayor participación relativa de los sitios vinculados a los servicios de construcción, educación y salud (gráfico 8).

Gráfico 8

Distribución de nuevos sitios web empresariales por rubro de actividad económica en Brasil, Colombia, México y Chile, entre marzo 2020 y octubre 2021
(En porcentajes)

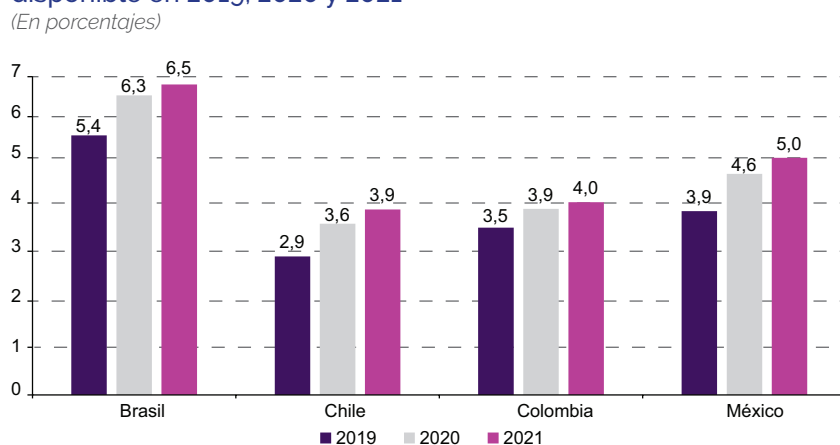


Fuente: CEPAL con datos de Dataprovider.

Asimismo, se observa un aumento en la cantidad de sitios web que ofrecen al menos una opción de pago en los cuatro países, lo que sugiere que las empresas aumentaron sus servicios transaccionales en línea, siendo congruente con la hipótesis del impulso a la digitalización producto del COVID-19 (véase gráfico 9).

Gráfico 9

Porcentaje de sitios web empresariales con al menos un método de pago disponible en 2019, 2020 y 2021
(En porcentajes)



Fuente: CEPAL con datos de Dataprovider.



A. Clasificación empresarial según la actividad en línea y definición de la economía de Internet

Como se indicó anteriormente, el presente estudio busca caracterizar a las empresas según su actividad en línea para dimensionar la economía de Internet, mediante el uso de macrodatos (*big data*), y en particular de técnicas de raspado web (*web scraping*) a través de la información pública disponible en Internet y su combinación con registros empresariales recopilados por las oficinas nacionales de estadísticas. La combinación de información web con registros oficiales permite tener un panorama más cercano a la realidad sobre las actividades que realizan las empresas formales en Internet, identificando grupos de empresas según su comportamiento en línea.

Este estudio tuvo como inspiración la prueba de concepto realizada por la Oficina Nacional de Estadística de los Países Bajos (CBS) que llevó a cabo un estudio sobre la “Medición de la economía de Internet en los Países Bajos: un análisis de Big Data” (Oostrom y otros, 2016). En este ejercicio, la información de los registros administrativos empresariales se complementó con los datos recopilados mediante técnicas de raspado web. El objetivo fue categorizar a las firmas en términos de su uso de Internet, con el fin de distinguir aquellas que hacen un uso más intensivo de esta tecnología como parte de su modelo de negocios y de esta manera estimar su contribución a la economía en general. Otro estudio, similar fue “Midiendo la economía digital del Reino Unido con Big Data” (Nathan y otros, 2016) realizado en el Reino Unido por la empresa Growth Intelligence y el Instituto Nacional de Investigaciones Económicas y Sociales, donde también se combinan diferentes fuentes de información como estadísticas oficiales, registros administrativos e información web para profundizar en el análisis de la economía digital.

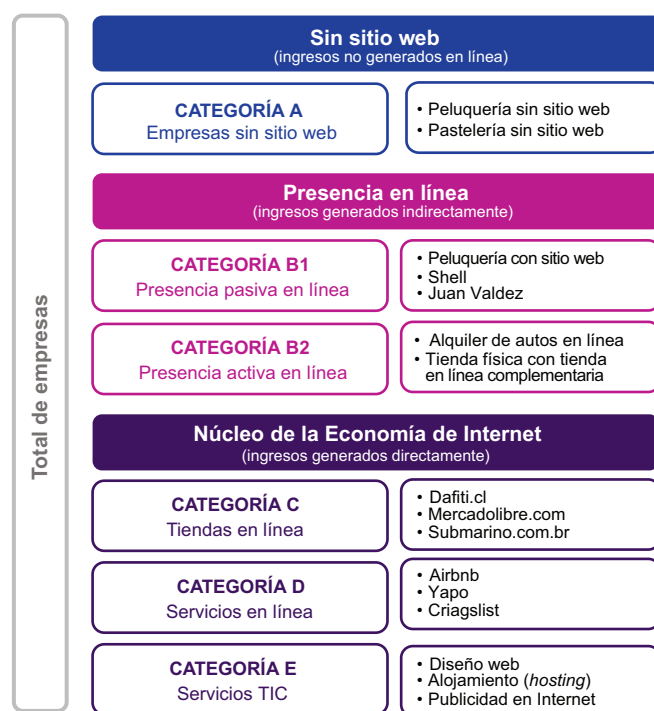
Los estudios citados encontraron conclusiones similares, que se pueden resumir en: i) la economía digital (o en su defecto la economía de Internet) es sustancialmente mayor a lo estimado por otras clasificaciones tradicionales; ii) las empresas que hacen un uso intensivo de Internet tienen ingresos medios superiores al resto de empresas; iii) la economía digital tiene una participación sustancialmente mayor en el empleo de lo que sugieren otras estimaciones convencionales.

Para fines del presente análisis se definirá a la economía de Internet como el conjunto de empresas que usan Internet en sus actividades empresariales y como parte de su modelo de negocio. A partir de esta aproximación se pueden organizar subcategorías según el nivel de ingresos generados en Internet⁶. En esta caracterización se distinguen tres categorías principales. La primera contiene a las empresas que no generan ningún ingreso en línea debido a su ausencia en la web (Categoría A), la segunda categoría se refiere a las empresas que generan ingresos en línea, pero de forma indirecta ya que su modelo de negocio no se basa exclusivamente en estos canales de comercialización, (Categoría B), y por último, las empresas que basan sus modelos de negocios en la web y por lo tanto, generan sus ingresos principalmente

⁶ Aquí adoptamos la definición de la economía de Internet desarrollada por la Oficina Nacional de Estadísticas de Países Bajos (CBS).

de sus actividades en línea y que representan el núcleo de la “economía de Internet” (Categorías C, D y E). En la categoría B, distinguimos además las categorías B1 y B2. La primera se refiere a las empresas que tienen una presencia en línea puramente pasiva, es decir, cuentan con un sitio web que contiene información sobre la empresa y sus productos o servicios, pero que no permite interactuar con ella. Por otro lado, los negocios incluidos en la Categoría B2 permiten algún tipo de interacción en línea, por ejemplo, suscribirse a un boletín informativo, registrarse para obtener una cuenta, hacer un pago o hacer una reserva. El núcleo de la economía de Internet comprende las empresas identificadas como tiendas en línea o plataformas de comercio-e que venden productos exclusivamente a través de sus sitios web (Categoría C), las empresas que ofrecen un servicio en línea (Categoría D), y finalmente, las empresas que proveen servicios y productos de TI. Cabe indicar, que las empresas de las categorías C y D incluyen modelos de negocio que básicamente no existirían sin Internet (véase diagrama 2).

Diagrama 2
Clasificación empresarial según su actividad en línea



Fuente: Adaptado de Ostrom L. y otros (2016).

Para recolectar la información de Internet se usaron técnicas de raspado web (*web scraping*) con una herramienta que permite indexar más de 450 millones de nombres de dominios en 50 países y que fue provista por Dataprovider⁷. Un aspecto importante para considerar en el presente estudio es que se utilizaron datos que están públicamente disponibles en línea. Estos datos no contemplan otra información transaccional en Internet y que se pueden producir en plataforma específicas de comercio electrónico (ej. MercadoLibre) y en redes sociales (ej. Facebook, Instagram). Esto constituye una limitante porque muchas micro, pequeñas y medianas empresas utilizan estos medios para realizar ventas en línea y promocionar sus productos y servicios.

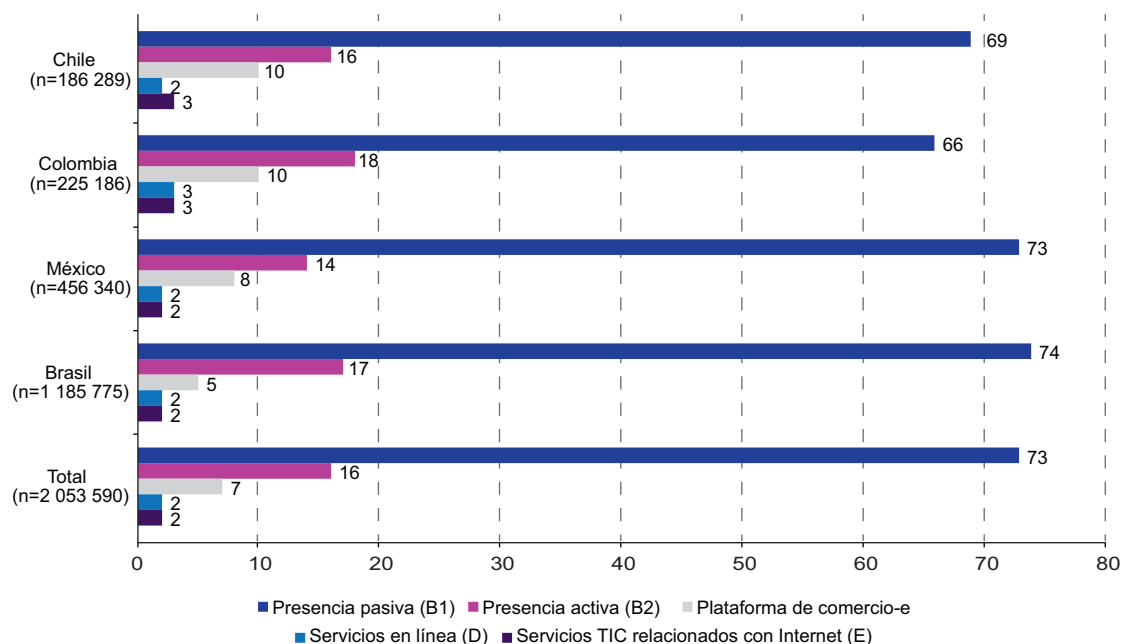
⁷ Una empresa con sede en Países Bajos especializada en ciencia de datos.

B. Composición de la economía de Internet a partir de datos web

Antes de examinar los resultados de la combinación de datos de sitios web con la información de los registros empresariales nacionales, se analizará primero la información capturada en la web para más de 2 millones de sitios web en los cuatro países. Por tal motivo, en esta parte del análisis la Categoría A no fue incluida. El Gráfico 10 muestra la distribución de los sitios web empresariales según su actividad en línea para los cuatro países a septiembre de 2021. Al observar la distribución de las cinco categorías (B1 – E), vemos que, en los cuatro países, la mayor parte de los sitios web de empresas, un poco más del 70%, tenía una presencia pasiva en línea (B1). Estas empresas no proporcionan ningún medio de interacción en línea con sus clientes, siendo solamente informativas. Los sitios web clasificados como con presencia activa en línea (B2) representaban entre el 14 % (México) y el 18 % (Colombia) del total de sitios web del respectivo país. Las empresas que poseen este tipo de sitios brindan algún medio para interactuar en línea con sus clientes. Finalmente, el núcleo de Internet (C-E) representaba entre el 9% (Brasil) y el 16 % (Colombia) de todos los sitios web comerciales y aproximadamente la mitad eran tiendas en línea. Cabe indicar, que la distribución es bastante similar entre los países analizados, aunque con algunas pequeñas diferencias.

Gráfico 10
Distribución de sitios web empresariales según tipo de actividad en línea por país, septiembre 2021

(En porcentajes del total de sitios web empresariales)

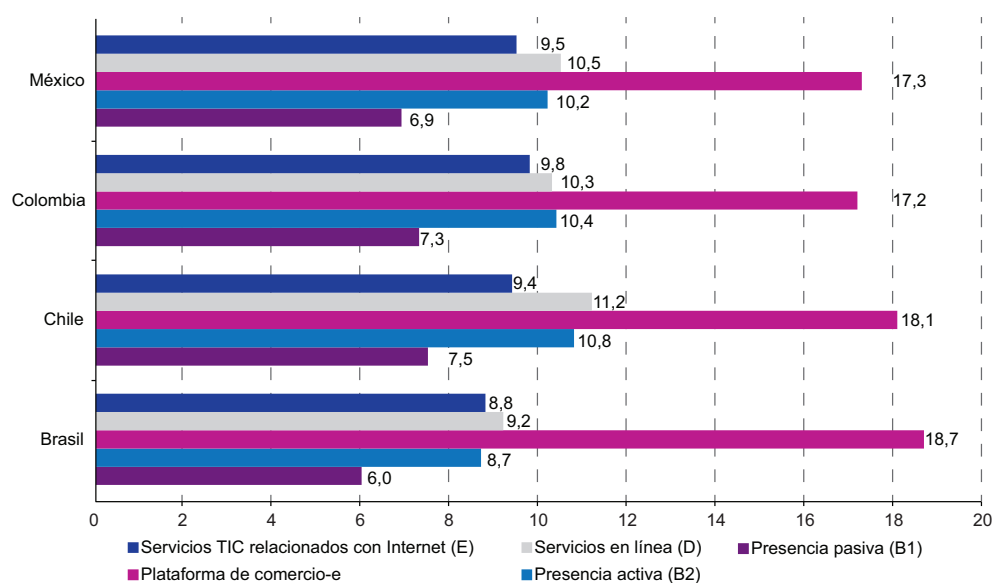


Fuente: CEPAL sobre la base de datos de Dataprovider.

A pesar de que la categoría B1 (presencia pasiva en línea) es, por mucho, la categoría más grande en los cuatro países, cuando se observa la actividad económica de todas las categorías, queda claro que las plataformas de comercio-e (categoría C) son las empresas con la huella económica más alta (véase gráfico 11). La huella económica es una métrica calculada por Dataprovider que da una estimación del volumen de

negocios del sitio web. La huella económica está en una escala logarítmica, por lo que los números más grandes pueden ser cientos o incluso miles de veces más grandes que los números más pequeños.

Gráfico 11
Huella económica de los sitios web empresariales según el tipo de actividad en línea, septiembre de 2021
(Índice)



Fuente: CEPAL sobre la base de datos de Dataprovider.

Este resultado podría explicar el hecho de que la cantidad de sitios web de comercio-e haya aumentado entre marzo de 2020 y septiembre de 2021, en contraposición al comportamiento del resto de las categorías. En ese periodo, luego del crecimiento explosivo producto del inicio de la pandemia en marzo de 2020, el total de sitios web empresariales disminuyó en Chile (-3%), México (-9%) y Brasil (-10%). Sólo aumentó 6% en Colombia. En los cuatro países, en ese período, la única categoría que registró crecimiento fue la de los sitios de comercio-e, con tasas significativas cercanas al 80% en los casos de Chile y Colombia, y de 55% en Brasil y México (véase gráfico 12). Esto significó la duplicación del peso de dicha categoría en el total de sitios web empresariales de los cuatro países hacia septiembre de 2021.

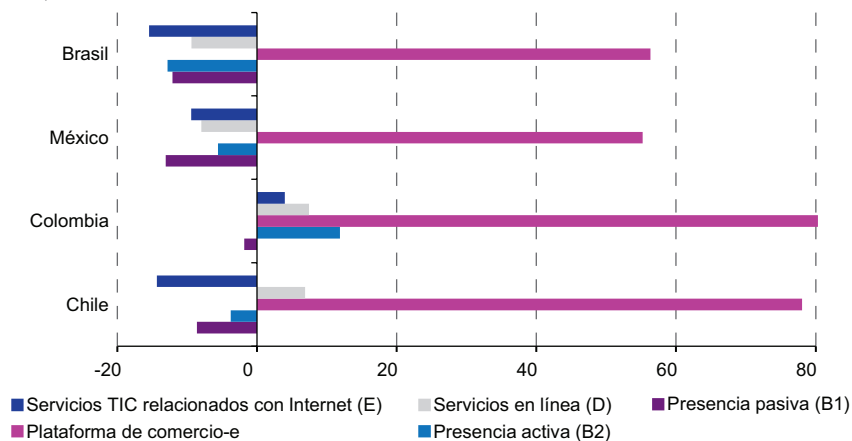
En un análisis de seguimiento, se revisaron específicamente aquellos sitios web que cambiaron de categoría de una presencia en línea pasiva a una presencia en línea más activa, observando también la cantidad de nuevos sitios web que surgieron por categoría. Este análisis reveló que la mayoría de los nuevos sitios web empresariales pertenecieron a la categoría B1: presencia en línea pasiva, seguidos de los nuevos sitios con una presencia activa en línea (B2) y de comercio-e (categoría C) (véase gráfico 13A). En tanto, entre los sitios web existentes, el mayor cambio fue el paso de una presencia en línea pasiva a una presencia activa, lo que significa que estas empresas agregaron funcionalidades a sus sitios web para brindarles a los clientes oportunidades para interactuar en línea. Asimismo, llama la atención que una proporción importante de sitios web empresariales hayan pasado de las categorías B1 (presencia pasiva en línea) y B2 (presencia activa en línea) a la categoría C (comercio-e), lo que indica que estas empresas comenzaron a vender productos en línea. Esto indicaría

que no existe una condición estructural que obstaculice a las empresas avanzar hacia una presencia web más sofisticada que denote una transformación de su modelo de negocios.

Gráfico 12

Variación de la cantidad de sitios web empresariales por tipo de actividad en línea entre marzo de 2020 y septiembre de 2021

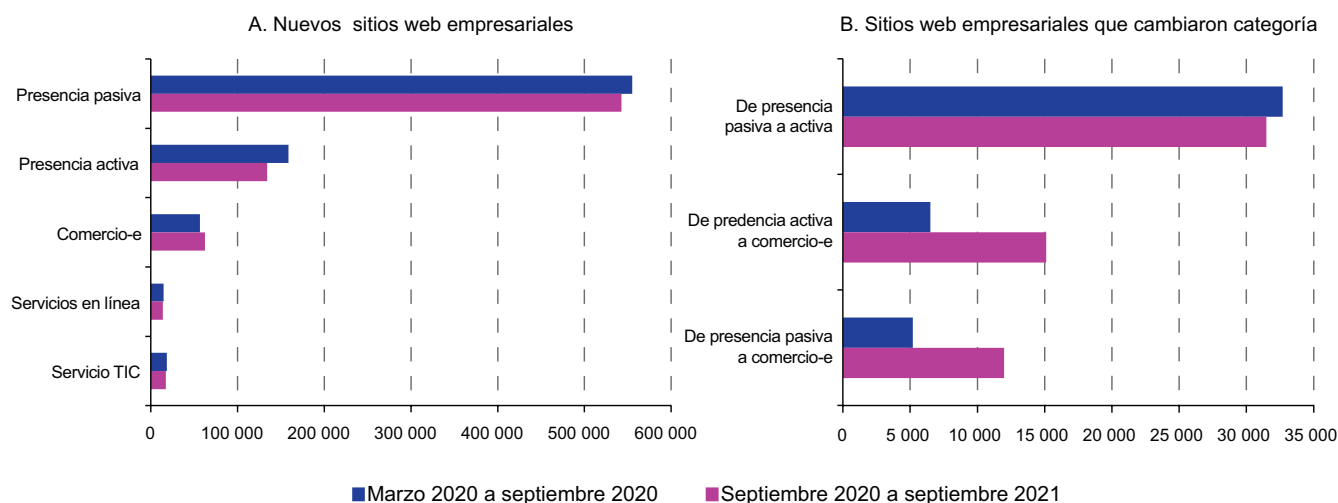
(En porcentajes)



Fuente: CEPAL sobre la base de datos de Dataprovider.

Gráfico 13

Sitios web empresariales nuevos y que cambiaron de categoría entre marzo y septiembre de 2020, y entre septiembre de 2020 y septiembre de 2021 en los cuatro países

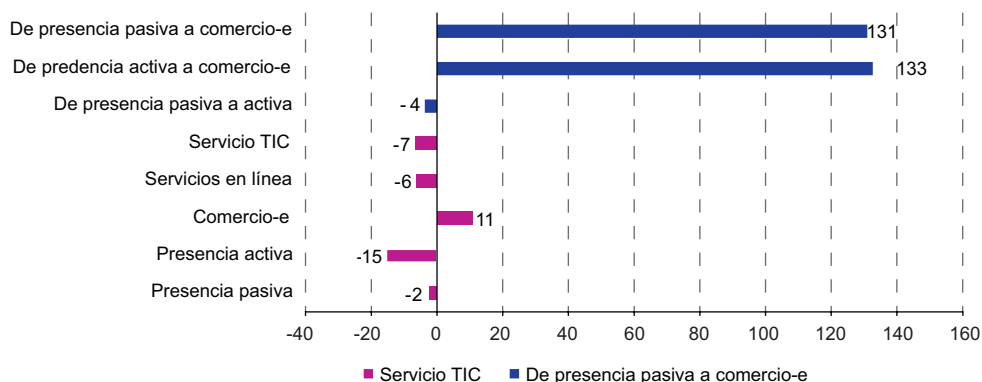


Fuente: CEPAL sobre la base de datos de Dataprovider.

Al comparar estos cambios, es decir el cambio de categoría de sitios existentes, como el apareamiento de nuevos, entre los periodos marzo y septiembre de 2020 versus septiembre de 2020 y septiembre 2021, se observa el constante avance hacia la modalidad de tienda en línea. Esta tendencia se mantuvo a pesar de una disminución en el surgimiento de nuevos sitios web (véase gráfico 14).

Gráfico 14

Variación en la cantidad de sitios web empresariales según categoría entre los periodos marzo-septiembre de 2020 y septiembre de 2020-septiembre de 2021 en los cuatro países



Fuente: CEPAL sobre la base de datos de Dataprovider.

C. La economía de Internet en Brasil, Chile, Colombia y México: resultados de la combinación de datos web y registros administrativos

A continuación, se presenta el análisis de la economía de Internet para Brasil, Chile, Colombia y México, con base en los resultados de la combinación de los datos de sitios web y los registros administrativos de las empresas. Este ejercicio fue posible gracias a la colaboración realizada con el Instituto Brasileiro de Geografía y Estadística (IBGE) y el Centro Regional de Estudios para el Desarrollo de la Sociedad de la Información (CETIC.br) de Brasil, el Instituto Nacional de Estadísticas de Chile (INE), el Departamento Administrativo Nacional de Estadística de Colombia (DANE), el Instituto Nacional de Estadística y Geografía de México (INEGI) y el NIC Chile. La vinculación más alta se logró en Brasil, donde aproximadamente el 9% de las empresas del registro de empresas se pudo relacionar con un sitio web correspondiente. Le siguieron Colombia (7%), Chile (3%) y México (1%). La mayoría de las empresas, más del 90%, se asignaron a la Categoría A (Empresas sin sitio web) de forma predeterminada. El cuadro 1 muestra el número total de empresas en el registro de cada país y las asignaciones de categorías correspondientes después de la vinculación.

Cuadro 1

Cantidad de empresas por categoría según actividad en línea para cada país, 2020

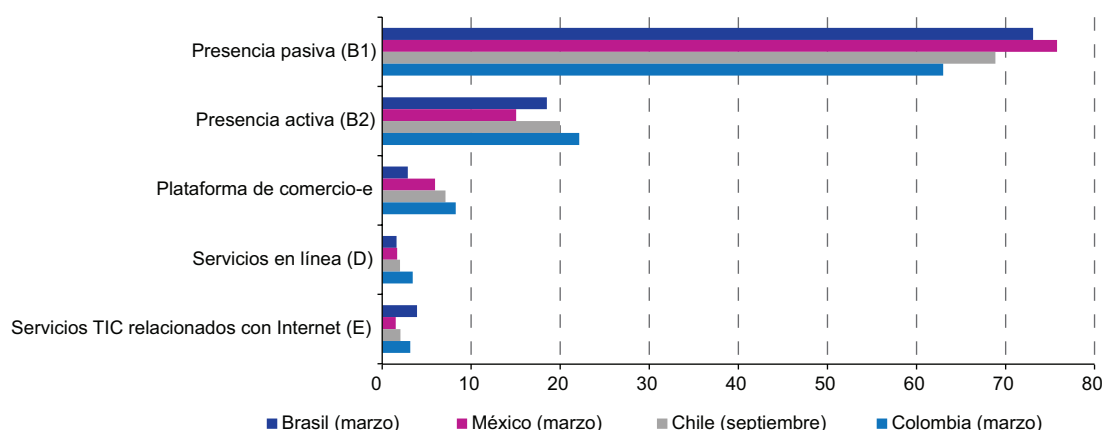
Categoría	Brasil (marzo)	México (marzo)	Chile (septiembre)	Colombia (marzo)
Empresa sin sitio web (A)	4 081 207	5 294 567	1 065 336	783 490
Empresas con presencia pasiva en línea (B1)	305 053	44 497	23 097	36 685
Empresas con presencia activa en línea (B2)	77 206	8 837	6 697	12 888
Tiendas en línea (C)	11 990	3 485	2 386	4 812
Servicios en línea (D)	6 757	993	672	1 995
Servicios TIC relacionados con Internet (E)	16 354	897	692	1 844
Total	4 498 567	5 353 276	1 098 880	841 714

Fuente: CEPAL con base en datos del IBGE (Instituto Brasileiro de Geografía y Estadística), CETIC.br (Centro Regional de Estudios para el Desarrollo de la Sociedad de la Información) de Brasil, INE (Instituto Nacional de Estadísticas) de Chile; DANE (Departamento Administrativo Nacional de Estadística) de Colombia; e INEGI (Instituto Nacional de Estadística y Geografía) de México y Dataprovider.

El gráfico 15 muestra la distribución de los sitios web vinculados por categoría según el tipo de presencia en línea por país. Se puede observar que la distribución de las categorías es similar en los cuatro países. Asimismo, como era previsible la mayoría de las empresas tienen una presencia pasiva en línea, entre el 63% y el 73% del total, y a esta categoría le siguen las empresas con una presencia activa. Las empresas del núcleo de Internet (categorías C, D y E) representan un número significativamente menor, pero con algunas diferencias por país. Por ejemplo, en Colombia, las tiendas en línea alcanzan cerca del 8% del total mientras que en Brasil representan el 3%.

Gráfico 15
Porcentaje de empresas vinculadas exitosamente con un sitio web
por categoría, 2020

(En porcentajes)

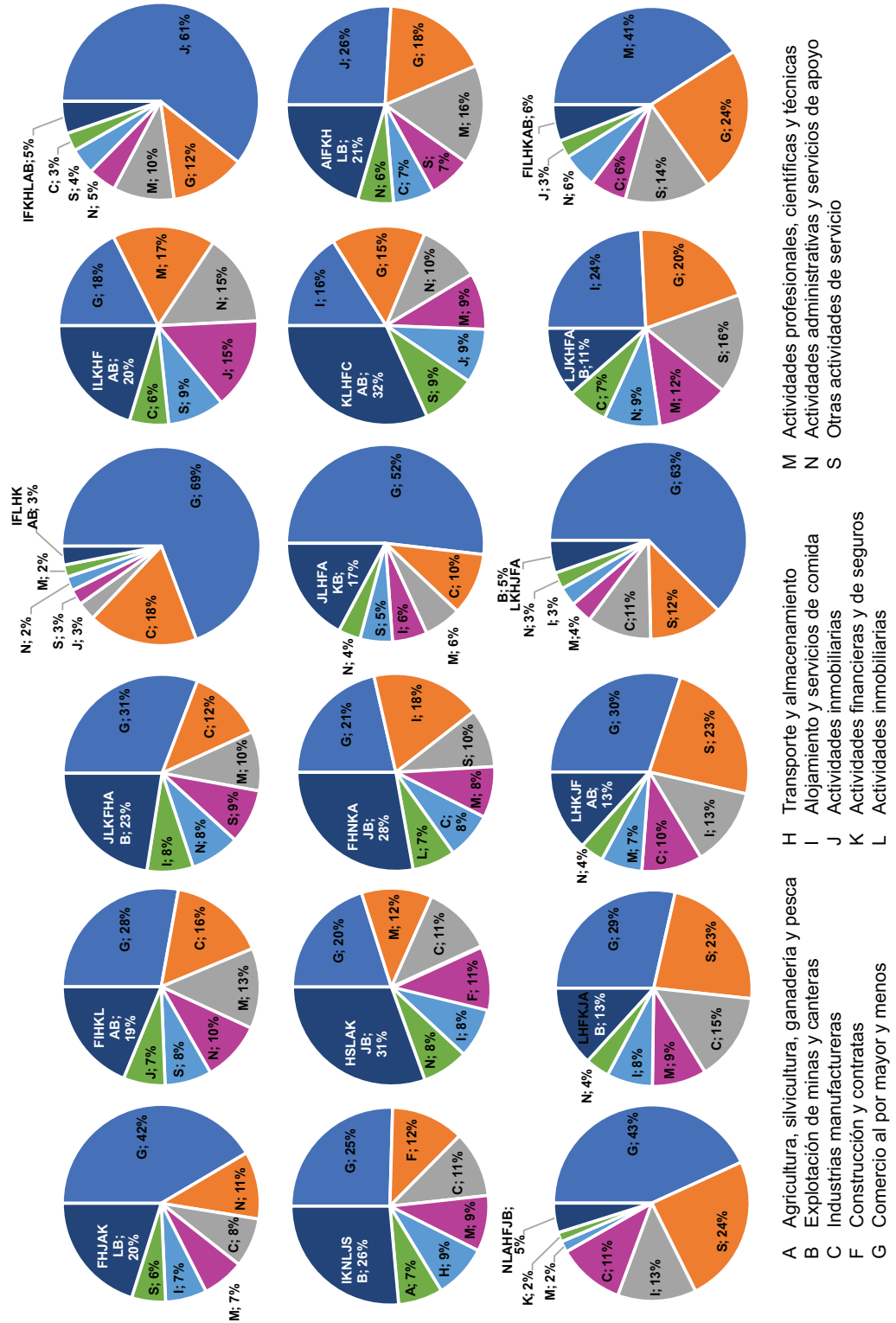


Fuente: CEPAL con base en el IBGE (Instituto Brasileiro de Geografia y Estadística), CETIC.br (Centro Regional de Estudios para el Desarrollo de la Sociedad de la Información) de Brasil, INE (Instituto Nacional de Estadísticas) de Chile, DANE (Departamento Administrativo Nacional de Estadística) de Colombia; e INEGI (Instituto Nacional de Estadística y Geografía) de México y Dataprovider.

A continuación, se examina la asignación de códigos CIU (Clasificación Industrial Internacional Uniforme) dentro de cada categoría para Brasil, Chile y México⁸ (véase gráfico 16). En casi todas las categorías exceptuando D (servicios en línea) y E (servicios TIC relacionados a Internet), el sector con mayor presencia es el sector de comercio (tiendas de comercio al por mayor y menor, sector G). Asimismo, este sector tiene una participación importante en la categoría de tiendas en línea (entre un 52% y 69%). En la categoría B1 (presencia pasiva), luego del sector de comercio, las principales actividades se vinculan con la manufactura (C), las actividades profesionales, científicas y técnicas (M) y otras actividades (S). En la categoría B2 (presencia activa), la distribución de actividades es similar a la categoría B1, pero con una presencia importante de los sectores de alojamiento y servicios de comida (I), lo que demuestra un nivel digitalización importante de estas empresas. En la categoría D (servicios en línea), luego del comercio, esta categoría se compone principalmente de las actividades profesionales, científicas y técnicas (M), las actividades administrativas (N), otros servicios (S), y el alojamiento y servicios de comida (I). En Brasil y Chile, los servicios de TIC e Internet (E) se vinculan con los sectores de información y comunicación como se espera de la definición de esta categoría. El hecho de que en México este sector no sea prominente probablemente refleja el bajo número de empresas que fueron vinculadas e indica que los datos no son representativos.

⁸ Los datos de Colombia no se incluyeron en esta representación ya que solo pudieron obtener datos de los siguientes SIC (C, G, H, I, J, K, M, N, P, R), ninguno de los cuales estaba representado en la Categoría C.

Gráfico 16
Participación de la industria (código SIC) por categoría de internet
 (En porcentajes)



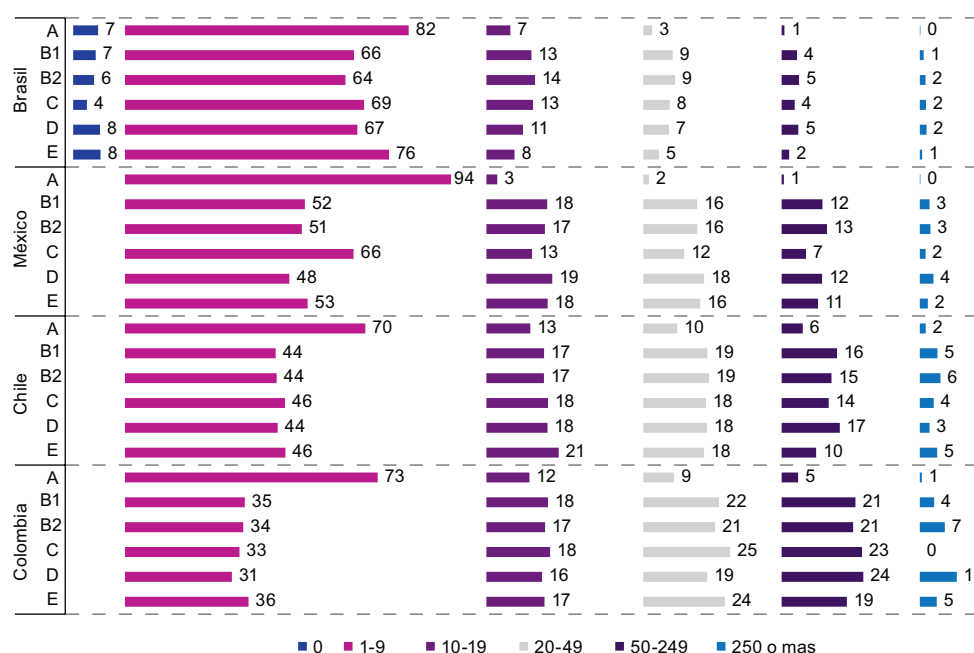
Fuente: CEPAL con base en el IBGE (Instituto Brasileiro de Geografia e Estatística), CETIC.br (Centro Regional de Estudios para el Desarrollo de la Sociedad de la Información) de Brasil, INE (Instituto Nacional de Estadísticas) de Chile; DANE (Departamento Administrativo Nacional de Estadística) de Colombia; e INEGI (Instituto Nacional de Estadística y Geografía) de México y Dataprovider.

En todos los países se pudo evaluar las categorías de la economía de Internet según el tamaño de la empresa. El gráfico 17 muestra la distribución del tamaño de la empresa por número de empleados para cada categoría. En los cuatro países, las empresas sin presencia web son en su mayoría pequeñas empresas con menos de 9 empleados y el porcentaje de grandes empresas (>250 empleados) sin sitio web es bajo. El resto de las categorías (B1-E) están distribuidas de forma uniforme en todos los países, excepto en Brasil donde la clasificación de 50 a 249 empleados es relativamente menor. Asimismo, en Brasil, las grandes empresas se distribuyen de manera uniforme en las categorías B1 a E. Llama la atención que, en Colombia, una proporción importante de grandes empresas se encuentran en la categoría de tienen en línea (D).

Gráfico 17

Participación de las empresas por categoría presencia en Internet según su tamaño (número de empleados)

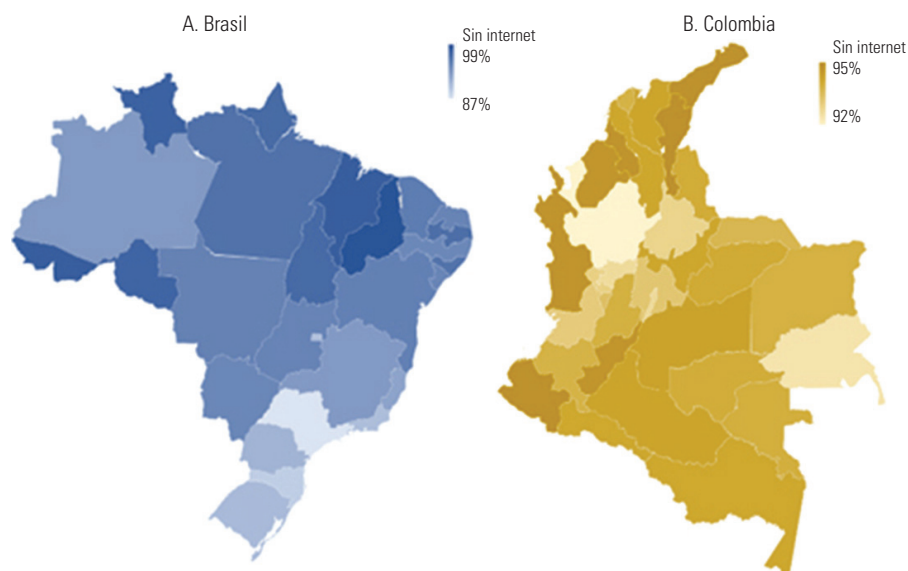
(En porcentajes)



Fuente: CEPAL con base en el IBGE (Instituto Brasileiro de Geografia y Estadística), CETIC.br (Centro Regional de Estudios para el Desarrollo de la Sociedad de la Información) de Brasil, INE (Instituto Nacional de Estadísticas) de Chile; DANE (Departamento Administrativo Nacional de Estadística) de Colombia, e INEGI (Instituto Nacional de Estadística y Geografía) de México y Dataprovider.

A través de los registros empresariales se puede analizar la distribución geográfica de las empresas con presencia web. Para Brasil este análisis pudo evidenciar una fuerte correlación entre la baja presencia en línea de las empresas y las regiones más desfavorecidas en términos de ingreso, como es el caso de la región nordeste de este país. En Colombia, los resultados muestran la misma correlación, pero al parecer también inciden otras variables como el conflicto armado, dado que las regiones con menor cantidad de empresas con presencia en línea están en los departamentos de El Cesar, La Guajira, Córdoba, Sucre, Choco y Nariño (véase mapa 1).

Mapa 1
Diferencias regionales de empresas sin sitio web para Brasil y Colombia
 (En porcentajes)



Fuente: CEPAL con base en el IBGE (Instituto Brasileiro de Geografia e Estatística), CETIC.br (Centro Regional de Estudios para el Desarrollo de la Sociedad de la Información) de Brasil, DANE (Departamento Administrativo Nacional de Estadística) de Colombia y Dataprovider.

Nota: Los límites y los nombres que figuran en este mapa no implican su apoyo o aceptación oficial por las Naciones Unidas.

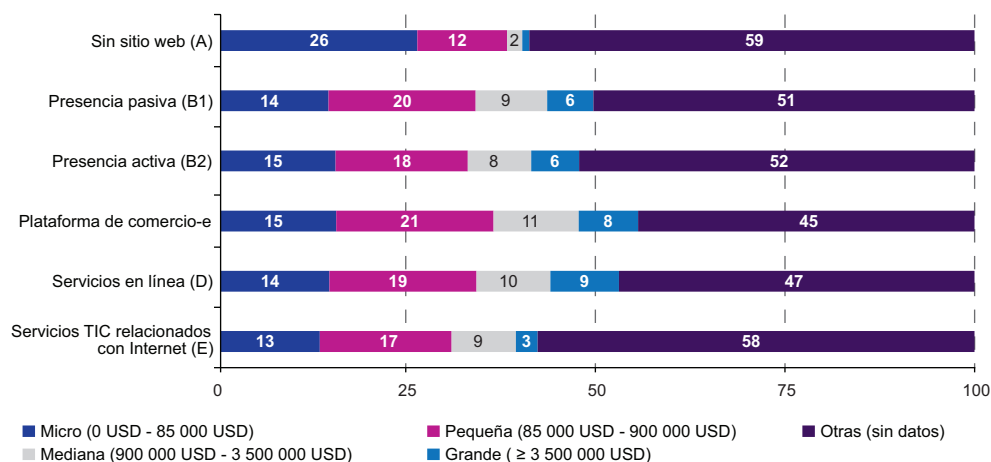
No todos los registros de empresas de los países disponen de la misma información. Los datos disponibles en Chile permitieron analizar el nivel de ventas anuales por categoría según el tipo de presencia en línea de las empresas, así como su orientación en cuanto a exportación/importación. El gráfico 18 muestra la participación de las empresas según su volumen de ventas anuales por categoría de presencia en línea. En marzo de 2020, las microempresas tenían aproximadamente la misma participación en todas las categorías, excepto en la categoría de A (sin presencia en línea), donde tenían preponderancia, lo que exhibe el bajo nivel de digitalización de este segmento. Contrariamente, las pequeñas, medianas y grandes empresas mostraron un nivel menor de participación en la categoría A al tener un grado mayor de digitalización. Asimismo, las categorías C (tiendas en línea) y D (servicios en línea) tienen una participación mayor de empresas medianas y grandes según su nivel de ventas. Esto demuestra la brecha digital entre los segmentos empresariales, pero también la oportunidad de generar mayores ingresos con base en la digitalización.

Al analizar la orientación de las empresas en materia de comercio exterior por su tipo de presencia en línea, se puede advertir claramente que, a marzo de 2020, las empresas chilenas sin presencia en línea operaban básicamente en el mercado local (véase gráfico 19). Otra observación es que la categoría C (tiendas en línea) tenía la mayor cantidad de empresas que se consideraban importadoras, a esta categoría le seguían la presencia en línea pasiva (B1) y activa (B2). Las otras categorías del núcleo de Internet (D y E) también registraban una proporción menor de empresas importadoras y exportadoras, lo que es lógico considerando que el tipo de actividad que realizan se vincula al área de servicios no transables.

Gráfico 18

Participación de las empresas según su volumen de ventas anuales por categoría de presencia en línea en Chile, marzo de 2020

(En porcentajes)

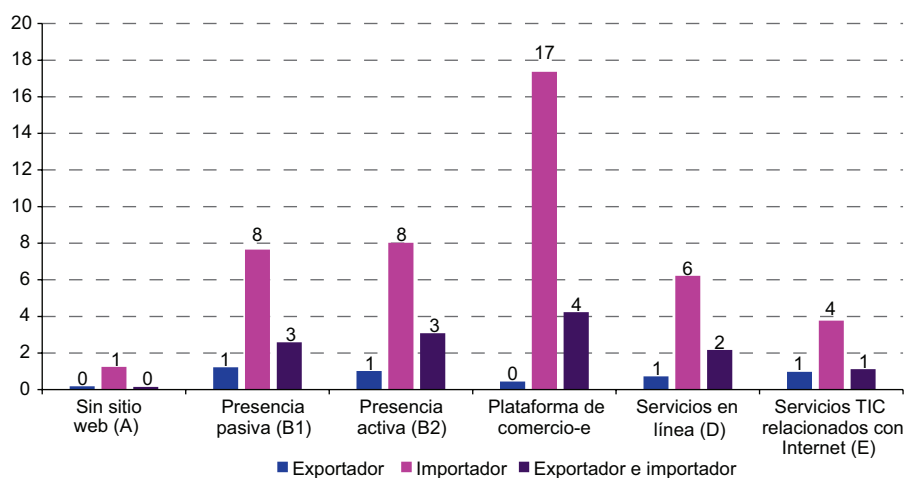


Fuente: CEPAL con base en INE (Instituto Nacional de Estadísticas) de Chile y Dataprovider.

Gráfico 19

Participación de empresas identificadas como exportadoras, importadoras o ambas, Chile, marzo de 2020

(En porcentajes)



Fuente: CEPAL con base en INE (Instituto Nacional de Estadísticas) de Chile y Dataprovider.

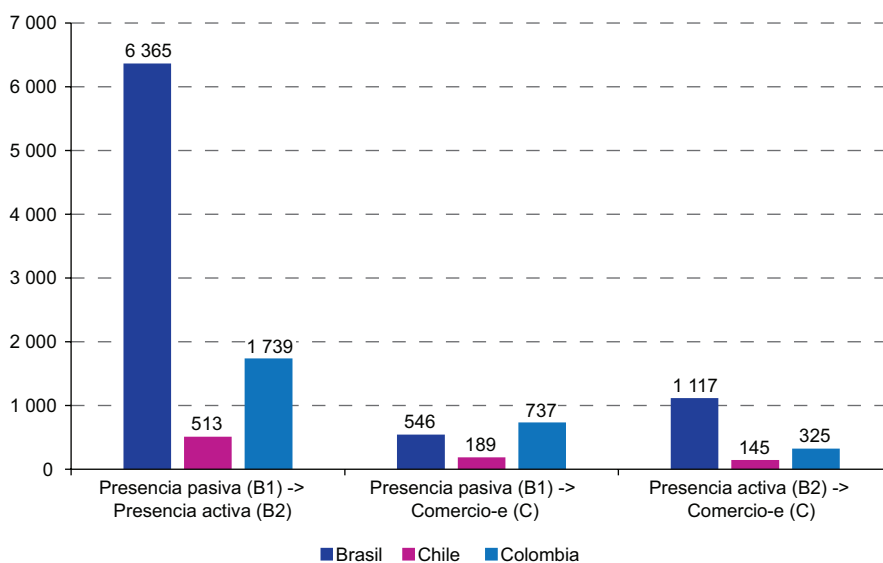
1. Efectos de la pandemia: seguimiento de los cambios entre marzo y septiembre de 2020

La pandemia de COVID-19 tuvo efectos importantes en el comportamiento de las empresas en línea. Como se indicó anteriormente, hubo un aumento en la cantidad de sitios web en general y una mayor sofisticación de las funciones de estos sitios para que los clientes interactúen en línea, así como también mayores ofertas de comercio electrónico. Los datos que se muestran a continuación analizan las empresas que pudieron ser vinculadas con éxito en los registros administrativos de los países del estudio, tanto en marzo como en septiembre de 2020, y que pasaron de una presencia en línea pasiva a una activa.

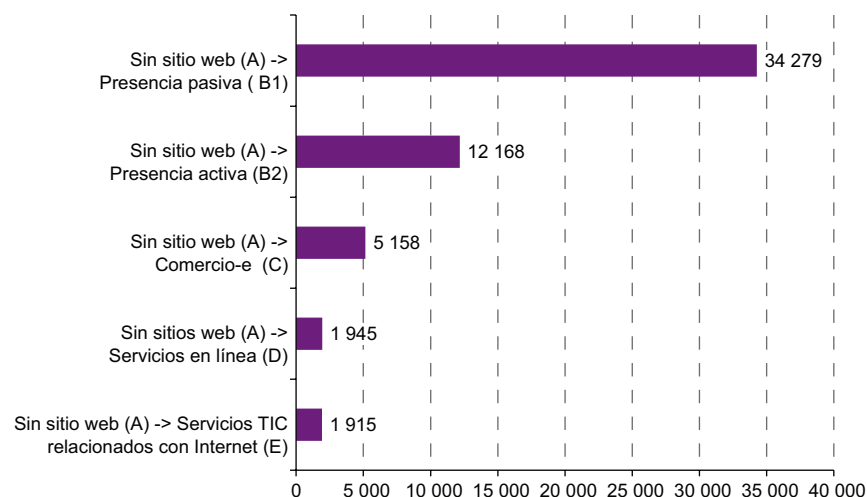
El gráfico 20 (a) muestra la cantidad de empresas que cambiaron de una presencia en línea pasiva a una activa o agregaron un componente de comercio electrónico en Brasil, Chile y Colombia (los datos para México no estaban disponibles). Un cambio de la categoría B1 a la categoría B2 significa que este sitio web ha agregado una característica adicional que inicialmente no estaba presente, por ejemplo, la posibilidad de realizar una reserva o suscribirse a un boletín informativo. Un cambio de B1 o B2 a C probablemente refleja la adición de actividades de comercio electrónico, por ejemplo, un sistema de carrito de compras o la aceptación de métodos de pago. El Gráfico 21 (b) muestra el cambio para Colombia de las empresas que no estaban en línea (Categoría A) en marzo a las diferentes categorías en línea en septiembre 2020. Estos datos demuestran no sólo el impacto de la pandemia en el comportamiento de la digitalización de las empresas, sino su capacidad para transitar a niveles más avanzados de uso de Internet.

Gráfico 20
Empresas que cambiaron su tipo de presencia en línea

A. Número de empresas que cambiaron de una categoría pasiva a una más activa



B. Número de empresas colombianas que pasaron de no tener presencia web a tener presencia en línea



Fuente: CEPAL con base en el IBGE (Instituto Brasileiro de Geografía y Estadística), CETIC.br (Centro Regional de Estudios para el Desarrollo de la Sociedad de la Información) de Brasil, INE (Instituto Nacional de Estadísticas) de Chile; DANE (Departamento Administrativo Nacional de Estadística) de Colombia y Dataprovider.

2. Consideraciones metodológicas en la combinación de datos web y registros empresariales

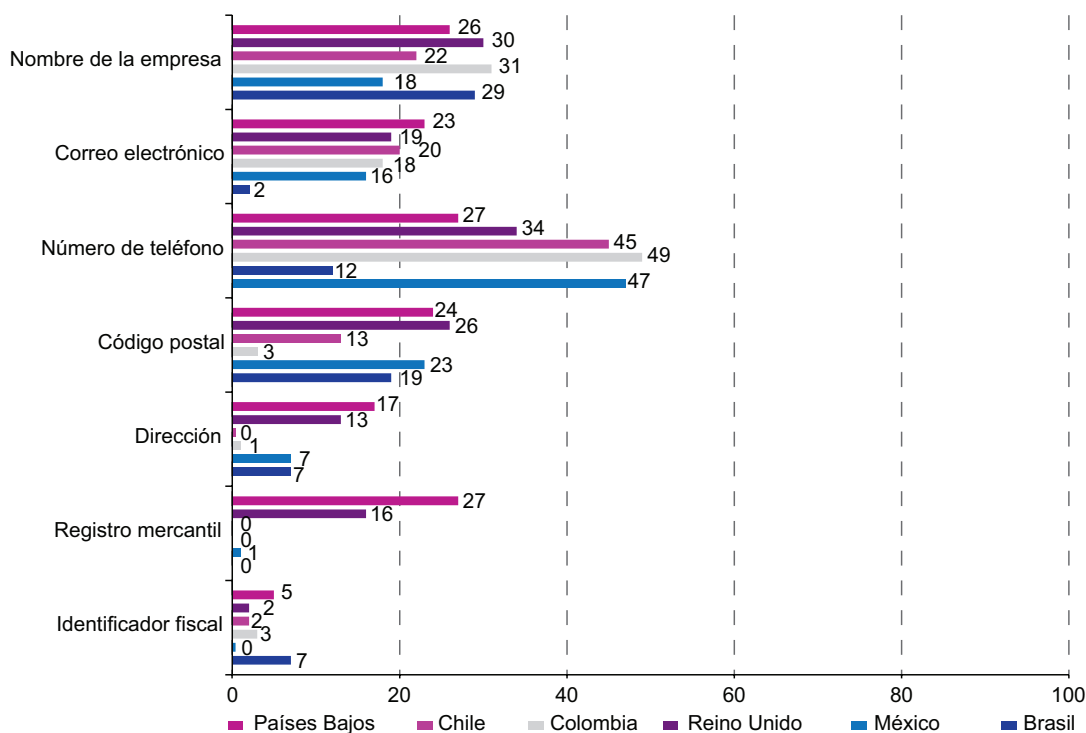
Uno de los principales desafíos en la fusión de la base de datos de sitios web con los registros nacionales de empresas de un país fue encontrar la información adecuada para hacer la vinculación de datos. Por ejemplo, en el estudio realizado por la Oficina Nacional de Estadística de los Países Bajos (CBS), el principal punto de combinación fue el número de registro comercial (CoC), un dato que la mayoría de las empresas de Países Bajos tienden a publicar en sus sitios web. Además, muchas empresas registran sus sitios web en el registro comercial, por lo que se pudo acceder a dos fuentes principales de datos para basar la combinación. Se utilizó información adicional como el correo electrónico y el teléfono solamente en los casos excepcionales, en los que no se contaba con las variables anteriores.

La disponibilidad de datos es un aspecto fundamental que diferencia el caso de Países Bajos con los cuatro países que participan en el presente estudio, dado que la cantidad de sitios web que brindan un número de registro comercial en estos países es muy baja. Por ello, se incluyeron otras variables para permitir una correspondencia entre el sitio web de una empresa y su inscripción en el registro comercial nacional. La información que se utilizó para realizar la vinculación incluye el nombre de la empresa, la dirección, el código postal, el número de teléfono con código de área y la dirección de correo electrónico. El gráfico 21 muestra la disponibilidad de esta información en los cuatro países del estudio y su comparación con los Países Bajos. Llama la atención que en menos del 1% de los sitios web se proporciona un número de registro comercial, en comparación con el 17% en los Países Bajos. Esta diferencia en la disponibilidad de información también es relevante para otras variables como el número de teléfono, no obstante, la disponibilidad es comparable en información más básica como en el nombre de la empresa.

En el gráfico 22 se examina la información disponible en los registros de empresas. Se observa, que en todos los países el número de empresas que tienen una URL registrada a nombre de su empresa va desde 0,7% en Brasil hasta un 5% en Chile. Del mismo modo, con la excepción de Brasil, las direcciones de correo electrónico están disponibles para menos del 13% de las empresas, lo que plantea grandes desafíos en la combinación de datos recolectados desde la web y los datos disponibles en los registros de empresas. La disponibilidad de direcciones es alta en los cuatro países, pero debido a la falta de información confiable sobre códigos postales en Chile y Colombia (información que se utiliza para verificar direcciones en la base de datos web), existe muy poca disponibilidad de datos en estos dos países.

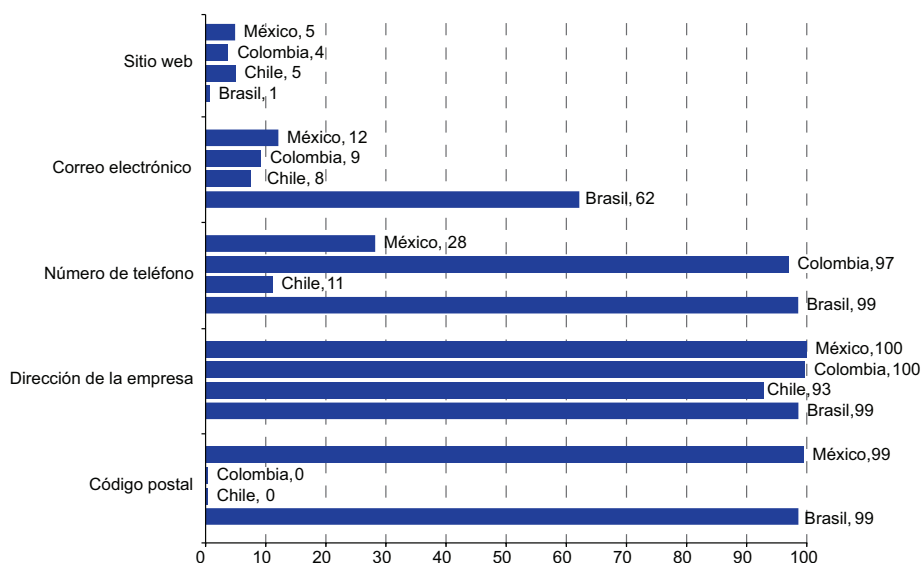
Las discrepancias indicadas anteriormente sobre la disponibilidad de información en los registros comerciales y la información web significaron una restricción al momento de combinar una mayor cantidad de información. En Colombia, la mayoría de las empresas se emparejaron utilizando números de teléfono (42.559) seguidos de la URL (14.972). Sólo se vincularon tres empresas en función de su número fiscal. En México, la mayoría de las empresas fueron combinadas mediante la URL (45,654), mientras que en otros casos (8,072) se pudo vincular la URL de la base de datos web con el nombre de la empresa. Solamente, 474 empresas se pudieron vincular utilizando el número de registro comercial en la base de datos web con el número de registro federal de contribuyente (RFC). En Brasil se realizó la combinación en dos etapas con la ayuda del NIC.br. En una primera combinación, se vinculó la base de datos web con el registro de NIC.br según el número comercial (CNPJ), luego, en una segunda etapa, se combinaron los números con el registro comercial (CEMPRE). Del total 377.001 (85,8%) de los negocios fusionados fueron identificados con éxito a través de su sitio web. El 14,4% restante se comparó mediante la identificación del CNPJ a través del sitio web y la información de correo electrónico de otras fuentes (por ejemplo, a través del Banco Central de Brasil – BACEN).

Gráfico 21
 Disponibilidad información en la base de datos web para seis variables
 utilizadas para la combinación (información a noviembre de 2021)
 (En porcentajes)



Fuente: CEPAL sobre la base de datos de Dataprovider.com.

Gráfico 22
 Disponibilidad de información dentro del registro de empresas
 de cada país para cinco variables utilizadas para la combinación
 (En porcentajes)



Fuente: CEPAL con base en el IBGE (Instituto Brasileiro de Geografia y Estadística), CETIC.br (Centro Regional de Estudios para el Desarrollo de la Sociedad de la Información) de Brasil, INE (Instituto Nacional de Estadísticas) de Chile; DANE (Departamento Administrativo Nacional de Estadística) de Colombia; e INEGI (Instituto Nacional de Estadística y Geografía) de México.

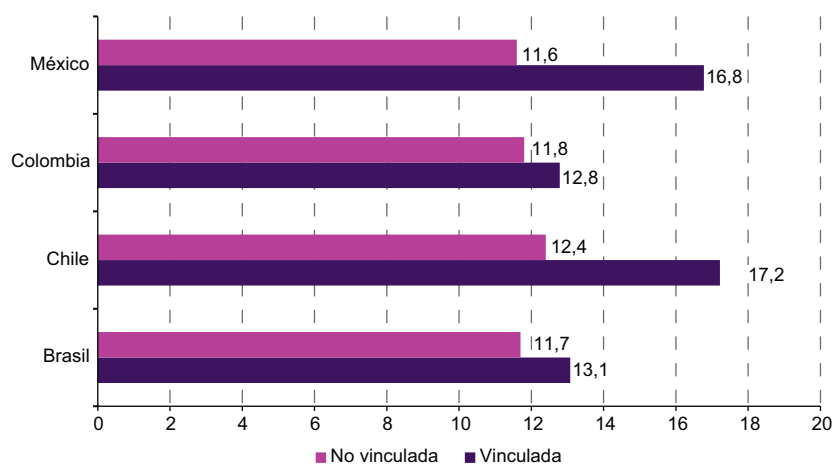
3. Características de los sitios web que fueron vinculados con el registro empresarial

Por último, en esta sección presentamos algunos datos preliminares que examinan las diferencias entre los sitios web de empresas que pudieron y no pudieron fusionarse con éxito con los registros nacionales de empresas. Esta sección sirve para comprender mejor la representatividad de los datos y para comprender mejor las características de las empresas que pudieron fusionarse. El gráfico 23 muestra la huella económica promedio de los sitios web que fueron y no fueron posible de vincularse para los cuatro países objetivo. Los sitios web que se fusionaron con éxito con empresas en el registro de empresas exhiben en general una huella económica más alta que los sitios web que no se relacionaron. También encontramos que las empresas fusionadas con éxito tienen una mayor presencia en algunas plataformas de redes sociales, incluidas Facebook y Twitter, mientras que hay poca diferencia con respecto a su presencia en otras plataformas de redes sociales, por ejemplo, Instagram (véase gráfico 24). En conjunto, los sitios web que pudieron fusionarse con los registros empresariales tienden a tener una mayor actividad en línea y, en general, un mayor impacto económico. Presumiblemente, estos sitios ofrecen más información sobre la empresa en sus sitios web, incluidos detalles de contacto importantes.

Gráfico 23

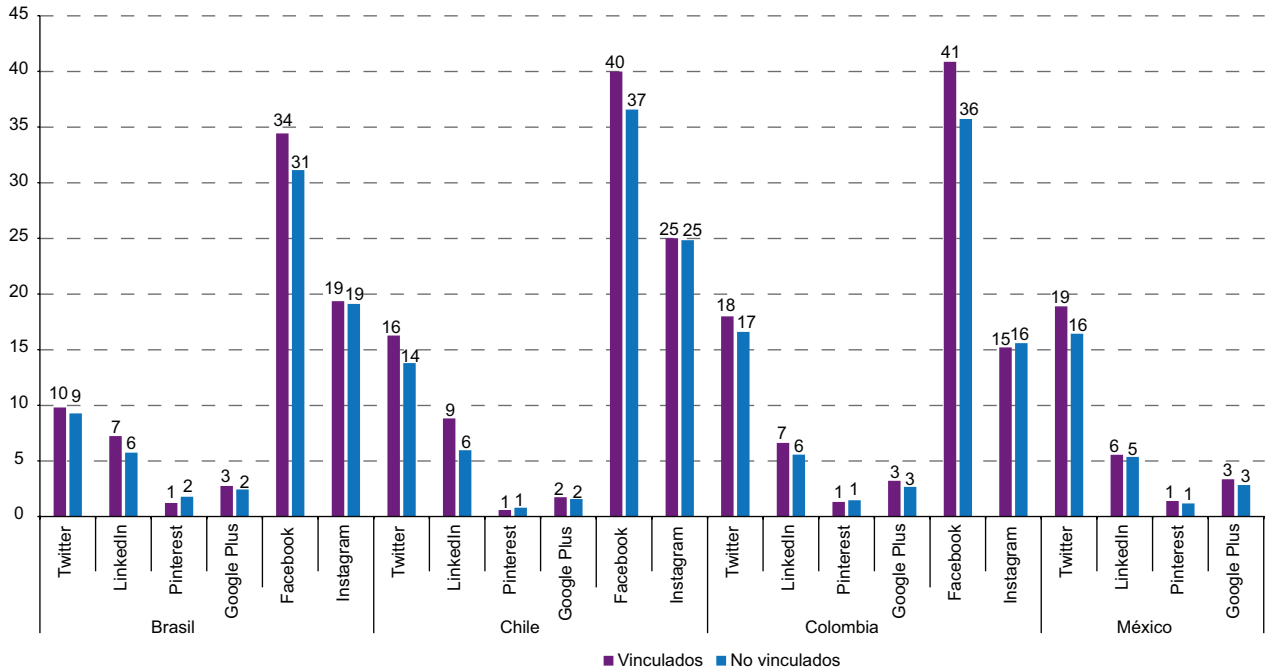
Huella económica promedio de los sitios web que se fusionaron y los que no se fusionaron en los registros comerciales nacionales

(Índice)



Fuente: CEPAL con base en Dataprovider.

Gráfico 24
Porcentaje de sitios web vinculados y no vinculados según
perfiles de redes sociales listados en el sitio web
(En porcentajes)



Fuente: CEPAL con base en el IBGE (Instituto Brasileiro de Geografía y Estadística), CETIC.br (Centro Regional de Estudios para el Desarrollo de la Sociedad de la Información) de Brasil, INE (Instituto Nacional de Estadísticas) de Chile; DANE (Departamento Administrativo Nacional de Estadística) de Colombia, e INEGI (Instituto Nacional de Estadística y Geografía) de México y Dataprovider.

Este estudio es una prueba de concepto que tuvo como objetivo principal caracterizar la economía de Internet en cuatro países latinoamericanos para comprender mejor cómo las empresas operan en línea. El estudio es un primer intento en la región de combinar fuentes de datos tradicionales con datos no estructurados provenientes de la web. Si bien este enfoque tiene el potencial de generar indicadores novedosos y revelar aspectos importantes de la economía digital, que de otra forma no sería posible, aún quedan varios desafíos importantes para aprovechar todo el potencial de estas nuevas fuentes de datos. A continuación, se resumen algunos retos de este tipo de herramientas y se realizan recomendaciones específicas para mejorar los resultados de ejercicios similares. Asimismo, se destacan los hallazgos más importantes sobre la economía de Internet.

Quizás el mayor desafío al que se enfrentó el proyecto fue vincular exitosamente la información disponible en un sitio web determinado con la información en los registros nacionales de empresas debido a la falta de datos superpuestos. En otros estudios similares realizados en otros países (por ejemplo, el realizado en Países Bajos), la mayoría de las empresas publican en su sitio web el número del registro comercial. Adicionalmente, el registro comercial en estos países también presenta información sobre los sitios web de las empresas. Ambas prácticas son menos comunes en los cuatro países que participaron en la investigación actual. Por lo tanto, sería útil que los registros nacionales solicitaran información sobre la presencia en línea de las empresas (sitio web y perfiles de redes sociales) en las encuestas nacionales. Además, alentar políticas para que las empresas publiquen información empresarial básica en su sitio web, como su registro comercial o fiscal, lo que ayudaría no solamente a identificar mejor a las empresas, sino que también ayudaría a proporcionar más transparencia hacia los consumidores. Un ejemplo, de este tipo políticas es el caso de Chile, que cuenta con un Reglamento de Comercio Electrónico del Ministerio de Economía, Fomento y Turismo, que desde 2021, dispone que las empresas que operan en plataformas de comercio electrónico entreguen información esencial, como el registro tributario, el domicilio legal y el correo electrónico en estas plataformas (Biblioteca del Congreso Nacional, 2021).


Entre los hallazgos principales del estudio se encuentra la ratificación que la mayoría de las empresas todavía mantiene una presencia pasiva en línea. No obstante, también se reveló que las empresas tienen la capacidad para transitar hacia estadios más avanzados en sus niveles de digitalización y uso de Internet. El impacto que tuvo la pandemia del COVID-19 en este cambio fue evidente. Asimismo, se pudo revelar las barreras que enfrentan las empresas más pequeñas al momento de establecer y administrar un sitio web. Cabe señalar, que en este caso el estudio no contempla el análisis de un componente importante de las actividades en línea de las empresas, especialmente las más pequeñas, que usan plataformas y redes sociales para mantener tener una presencia en línea. En futuros trabajos, se deberá intentar integrar y capturar estas actividades para comprender mejor esta dinámica empresarial.

Si bien no se pudo acceder a datos para todos los países en relación con los ingresos que generan las empresas, los datos disponibles (Chile), exhibieron algunos hallazgos, por ejemplo, las categorías que pertenecen al núcleo de Internet como las tiendas en línea y los servicios en línea tienden a generar más ingresos que las

empresas con un nivel menor grado de digitalización. Igualmente, el estudio demostró, para aquellos países donde se pudieron acceder a los datos (Brasil y Colombia), la brecha geográfica de la presencia en línea de las empresas, y su posible correlación con los niveles de ingreso y otros problemas como el despliegue de infraestructura. En relación con la mirada sectorial, también se pudo confirmar, como era de esperar, que los sectores de servicios tienen una mayor participación en el núcleo de Internet, pero que también existe una participación interesante de la manufactura en estas categorías.

Finalmente, a pesar de algunos desafíos importantes, este proyecto ha creado un conocimiento valioso, allanando el camino para futuras investigaciones donde se aprovechen la cantidad de datos que produce la huella digital web y que es cada vez mayor. Particularmente, se pudo demostrar su utilidad para analizar el comportamiento empresarial y la economía de Internet. Asimismo, se pudo comprobar algunos de los profundos efectos que tuvo la pandemia del COVID-19 en las actividades comerciales en línea. Sin duda, las empresas se ven cada vez más exigidas en digitalizar sus actividades, una tendencia que probablemente solo se acelere en los próximos años. Por este motivo, es crucial que los gobiernos cuenten con información para entender esta dinámica a fin de generar políticas que faciliten e impulsen la transformación digital productiva.

- Bailey, M. y otros (2012), "The Menlo Report", *IEEE Security & Privacy*, vol. 10, No. 2, marzo.
- CEPAL (2022), *Un camino digital para el desarrollo sostenible de América Latina y el Caribe*, Santiago, Comisión Económica para América Latina y el Caribe, noviembre.
- ___ (2021), "La Inversión Extranjera Directa en América Latina y el Caribe 2021", *CEPAL* (LC/PUB.2021/8-P).
- ___ (2020), *Análisis de la huella digital en América Latina y el Caribe: enseñanzas extraídas del uso de macrodatos (big data) para evaluar la economía digital*, Proyecto Big Data para la Medición de la Economía Digital, Cuenta de las Naciones Unidas para el Desarrollo, Santiago, Comisión Económica para América Latina y el Caribe, julio.
- ___ (2018), "Datos, algoritmos y políticas: la redefinición del mundo digital", CEPAL, Naciones Unidas, Santiago, abril.
- CEPAL & Adenauer, F. K. (2021), *Recuperación económica tras la pandemia COVID-19: empoderar a América Latina y el Caribe para un mejor aprovechamiento del comercio electrónico y digital*, Santiago, CEPAL, Naciones Unidas, octubre.
- CEPAL, CAF, & Digital policy law (2020), *Las oportunidades de la digitalización en América Latina frente al COVID-19*, Santiago, Comisión Económica para América Latina y el Caribe, abril.
- Davahli, M. R. y otros (2020), "Identification and Prediction of Human Behavior through Mining of Unstructured Textual Data", *Symmetry*, vol. 12, 19 de noviembre.
- European Commission (2014), "COMmission expert group on taxation of the digital economy", Paris, European Commission, mayo.
- Layne, D. (2001), *3D Data Management: Controlling Data Volume, Velocity, and Variety*, META Group, enero.
- Markovikj, D. y otros (2013), "Mining Facebook Data for Predictive Personality Modeling", *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 7, No. 2.
- Nacional, B. del C. (2021), "Biblioteca del Congreso Nacional | Ley Chile", www.bcn.cl/leychile, [en línea] <<https://www.bcn.cl/leychile>> [fecha de consulta: 20 de abril de 2023].
- Nathan, M. y otros (2016), "Measuring The UK's Digital Economy With Big Data".
- OECD (2020), "A Roadmap Toward A Common Framework For Measuring The Digital Economy", Saudi Arabia, Report for the G20 Digital Economy Task Force, enero.
- ___ (2011), *OECD Guide to Measuring the Information Society 2011*, Paris, Organisation for Economic Cooperation and Development, agosto.
- Ostrom, L. y otros (2016), "Measuring the internet economy in The Netherlands: a big data analysis".
- UNCTAD (2022), "Manual para la producción de estadísticas sobre la economía digital", Ginebra, Conferencia de las Naciones Unidas sobre Comercio y Desarrollo.
- Ward, J. S. & Barker, A. (2013), "Undefined By Data: A Survey of Big Data Definitions". Publicación de las Naciones Unidas.



La adopción de tecnologías digitales es un instrumento esencial para reducir las brechas de productividad entre los países de América Latina y el Caribe y los países más desarrollados, generar nuevas fuentes de crecimiento y la creación de empleos de calidad. Para aprovechar estas tecnologías se necesitan políticas basadas en datos empíricos que permitan encauzar el cambio tecnológico, aprovechando sus oportunidades y reduciendo sus riesgos.

En esta publicación se presenta un ejercicio exploratorio realizado en el Brasil, Chile, Colombia y México en el que se combinan fuentes alternativas de información, particularmente datos extraídos de la web, con fuentes oficiales de estadísticas para medir la actividad empresarial en línea. Esta metodología permite clasificar a las empresas según el uso que hacen de Internet, sin necesidad de restringirse a clasificaciones industriales tradicionales, lo que genera una nueva caracterización: la economía de Internet. Asimismo, el estudio permite experimentar sobre las posibilidades que ofrece el uso de técnicas y herramientas de grandes datos para mejorar la comprensión de la transformación digital y sirve como base para futuras investigaciones.

